Organisation for Economic Co-operation and Development

ENV/CBC/MONO(2025)12

Unclassified

English - Or. English

2 October 2025

ENVIRONMENT DIRECTORATE
CHEMICALS AND BIOTECHNOLOGY COMMITTEE

Cancels & replaces the same document of 30 September 2025

Guiding principles for mixture threshold derivation from effect biomarkers

Series on Testing and Assessment, No 414

JT03572109

Please cite this publication as:

OECD (2025), *Guiding principles for mixture threshold derivation from effect biomarkers,* OECD Series on Testing and Assessment, No 414, OECD Environment, Health and Safety, Paris, https://one.oecd.org/official-document/ENV/CBC/MONO(2025)12/en

Contact us

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

E-mail: ehscont@oecd.org

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 countries in North and South America, Europe and the Asia and Pacific region, as well as the European Union, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several Partner countries and from interested international organisations attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials;** and **Adverse Outcome Pathways.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site.

Guiding principles for mixture threshold derivation from effect biomarkers

Table of contents

Tables

Figures

Foreword

Currently available assessment approaches for cumulative risk of chemical mixtures can only be applicable to a small number of substances present at workplaces and in the environment. We cannot anticipate a significant change of this situation in the near future due to extensive data need for cumulative risk assessments. Presently, effect biomarkers are the most direct option to address the risk of known and unknown mixtures in an integrative way. Traditional occupational health risk assessments often rely on external exposure measurements, such as air monitoring, which may not fully capture the complexities of workplace exposures. Human biomonitoring is used to measure internal exposures or effects in exposed individuals or groups from all potential routes of exposure (i.e., inhalation, oral, and dermal). Exposure to mixtures in the workplace and environment is the most common chemical exposure scenario in our daily lives. However, methods for assessing the risks and for setting mixture threshold limits to avoid adverse effects lack global harmonization. Monitoring of effect biomarkers can support regulatory risk assessment in multiple ways.

Note: terms marked with an "" are defined in the glossary and abbreviations section at first time mentioning.*

An effect biomarker* indicates a stressor-induced biological effect which can be associated with a disease and can be interpreted as a potential predictor of a downstream effect i.e. measuring a key event in a Mode of Action (MoA) or Adverse Outcome Pathway (AOP). Thus, biomarkers can provide an integrated measure of the response to relevant stressors by all routes of known and unknown exposures. However, effect biomarker responses are usually not straightforward to interpret regarding their predictive value to indicate adverse effects. A systematic understanding of the relevance of effect biomarker data will enhance the protection of workers and/or ecosystems, if used under appropriate ethical and regulatory frameworks. Therefore, harmonized guidance for assessing effect biomarkers and their application to risk assessments are needed.

The guiding principles proposed in this document describe the key concepts for the derivation and interpretation of mixture thresholds* for selected effect biomarkers for use in occupational or ecological risk assessments. The aim of these guiding principles is to present a harmonized assessment approach which will save resources and promote consistency across regulatory agencies at national and international levels.

The development of this document was a joint activity of the Organisation for Economic Co-operation and Development (OECD) Working Party on Exposure Assessment & Working Party on Hazard Assessment (WPEA & WPHA) in collaboration with more than 90 experts from 25 countries and other stakeholders (see chapter 8 project participation). The activity was started in October 2022 and the development of this guiding principles document was co-led by Robert Pasanen-Kase (SECO*, CH) as coordinator, Maryam Zare-Jeddi (BIAC*), Nancy B. Hopf (Unisanté, CH), Susana Viegas (ENSP*/UNL, PT), Dan Villeneuve (US-EPA*, US), Martin Wilks and Rex FitzGerald (University of Basel, CH), Radu Corneliu Duca (LNS*, LU) and the OECD Secretariat. The document was drafted in close collaboration with experts providing input on different aspects of human and environmental effect-biomonitoring including Bernice Scholten, (TNO* , NL), Eszter Simon (FOEN* , CH), Devika Poddalgoda (Health Canada, CAN), Anna Bal Price

(JRC*, EU); Vicente Mustieles, Antonio Hernandez-Jerez (University of Granada, ES), Christoph van Thriel (IFADO* , DE), Stefano Bonassi (IRCCS*, San Raffaele Roma, IT), Michael Fenech (University of South Australia, AUS), Sophie Ndaw (INRS*, FR), Christina Pieper (German Federal Institute for Risk Assessment, DE), Lucian Farcal, Alicia Paini (EFSA*, EU).

The initial draft guidance document was reviewed in 2025 by expert group (see chapter 8) and WPEA & WPHA members and was commented by eleven experts from six different organisations / institutes / companies and was finalized. This adopted biomonitoring guiding principles document is published under the responsibility of the Chemical and Biotechnology Committee of the OECD.

Glossary and Abbreviations

ACGIH:= American Conference of Governmental Industrial Hygienists

ADME:= Absorption, Distribution, Metabolism, Excretion

AO:= Adverse Outcome in an AOP.

AOP:= Adverse Outcome Pathway.

Brief: An AOP describes a chain of events at different levels of biological organisation that causally connects a molecular initiating event to an adverse health outcome (Beronius et al. 2021). Expanded: An AOP consists of a single sequence of key events connecting an initial upstream molecular initiating event (MIE) to an adverse outcome (AO) such as a human disease. Key events must be measurable, necessary, and consistent, but may not be sufficient to trigger an adverse outcome. An AOP describes toxicodynamic key events, not toxicokinetic key events prior to the MIE, and is thus chemical-agnostic. If chemical-specific toxicokinetic key events are included, the AOP is a mode of action (MoA) (Edwards et al. 2016).

BDNF:= Brain Derived Neurotrophic Factor

BEQ:= Bioanalytical Equivalent Concentration

BIAC:= Business and Industry Advisory Committee to the OECD advocates for policies at OECD for the private sector

DNT:= Developmental Neurotoxicity

E2-EQ:= 17-beta estradiol (E2) equivalents

EBM:= Effect-Based Methods

EBT, **P-EBT, R-EBT,T-EBT:=** Effect-Based-Trigger value, can be used as mixture threshold to assess environmental exposures and is divided in the subtypes Provisional (P)-, Reference (R)-, Technical (T)-EBT.

ED:= Endocrine Disruption

Effect biomarker:= A measurable biochemical, physiological, behavioral or other alteration within an organism that, depending upon the magnitude, can be recognized as associated with an established or possible health impairment or disease (WHO 1993).

Exposure biomarker:= chemical substances, substance groups or their metabolites in biological matrixes

EFSA:= European Food Safety Authority

ENSP:= Escola Nacional de Saúde Pública

EU:= European Union

EQS:= Environmental Quality Standard

FOEN:= The Federal Office for the Environment

GHS:= United Nations Globally Harmonized System of Classification and Labelling of Chemicals

HBM:= Human Biomonitoring

IARC:= International Agency for Research on Cancer

ICCS:= Istituto di Ricovero e Cura a Carattere Scientifico

ICAPO:= International Council on Animal Protection in OECD Programmes seeks to ensure the widest possible integration of alternative methods.

IFADO := Leibniz Institut für Arbeitsforschung an der TU Dortmund

INRS:= Institut National de Recherche et de Sécurité

JRC:= Joint Research Centre of the European Commission

KE:= Key Event in an Adverse Outcome Pathway (AOP)

L-CBMN:= Lymphocyte Cytokinesis-Block Micronucleus assay

LOD:= Limit of Detection

LOQ:= Limit of Quantification

LNS:= Laboratoire Nationale Santé

MIE:= Molecular Initiating Event in an Adverse Outcome Pathway (AOP)

Mixture threshold:= Specific level of change in an effect biomarker response that indicates an exposure to chemical mixtures leading to an adverse health risk in human or environmental organisms. These thresholds indicate the magnitude of biomarker response that would prompt regulatory actions aimed at protecting human and environmental health. The term "mixture threshold" is interchangeable with "effect based trigger values" used in ecological risk assessment.

MN:= Micronucleus

MoA:= Mode of Action. A MoA consists of sequential key events leading from exposure to a substance to a final specific toxic outcome. Key events must be measurable, necessary, and consistent, but may not be sufficient to trigger a toxic outcome. A MoA describes both toxicokinetic and toxicodynamic key events, and is thus chemical-specific (Carmichael et al. 2011).

NfL:= Neurofilament Light Chain

Ng:= Neurogranin, also NGRN

NT:= Neurotoxicity

OBEL, POBEL, ROBEL, TOBEL:= Occupational Biomonitoring Effect Level can be used as mixture threshold to assess occupational exposure effects and is divided in the subtypes Provisional (P)-, Reference (R)-, Technical (T)-OBEL.

OBL:= Occupational Biomonitoring Level

OEL:= Occupational Exposure Limit

OHP:= Occupational Health or Hygiene Professional

PAH:= Polycyclic Aromatic Hydrocarbon

PCB:= Polychlorinated biphenyls

PPE:= Personal Protective Equipment

RMM:= Risk Management Measures

SECO:= State Secretariat for Economic Affairs

SEG:= Similar Exposure Group

SHBG:= Sex Hormone-Binding Globulin

TNO:= Netherlands Organisation for Applied Scientific Research

TT:= Testosterone

UNL:= University of Lisbon

US-EPA:= United States of America Environmental Protection Agency

VTG:= Vitellogenin

WHO:= World Health Organization

1 Rationale for using effect biomarkers in human health and environmental context

There is an almost infinite number of chemical combinations to which humans and the environment can be exposed. Most government authorities regulate single chemicals, although humans and the environment are exposed to chemical mixtures in complex and in often not measurable combinations. Measuring effect biomarkers can be used as an approach to evaluate exposures to chemical mixtures, even in circumstances where the mixtures are only partially defined.

Although it is widely acknowledged that multiple regulated and unregulated substances are simultaneously present in the environment and within the organism, current legislation is still predominantly focused on individual substances and is typically enforced based on limit value analysis of a limited subset of substances. Most available human and environmental biomonitoring studies have focused on a limited number of substances that are well-known for their toxicity and the availability of analytical techniques. This limits the scope of mixture assessment strategies to only a small number of substances. However, these limitations can be overcome by utilizing effect biomarkers, which can directly assess effects from exposures of known and unknown chemical mixtures. Effect biomarkers facilitate risk assessment of complex mixtures, especially if used together with exposure biomonitoring. Both exposure and effect biomarkers have their own strengths and limitations.

An effect biomarker indicates a stressor-induced biological effect which can be associated with a disease and can be interpreted as a potential predictor of a downstream effect i.e. measuring a key event in a Mode of Action or Adverse Outcome Pathway. Thus, biomarkers can provide an integrated measure of the response to relevant stressors by all routes of exposure. Effect biomarkers linked to a well-characterized MoA or AOP can be used as early indicator for assessing the risk of adverse health effects from exposure to chemical mixtures. An ideal effect biomarker should be sensitive, specific, robust, predictive of an adverse outcome or downstream event, and preferably non-invasive. Biomarkers of exposure* are chemical substances themselves (parent compounds), substance groups, or their breakdown products in the body (metabolites), quantified in biological matrices. Biomarkers of exposure are complementary to effect biomarkers.

Although, biomonitoring of effect biomarkers is the most direct option for measuring effects from both known and unknown components of a chemical mixture in an integrative way, they are used less often than exposure biomarkers. Possible reasons include the lack of specificity with regard to specific chemical exposures, and uncertainties in causal relationships to apical endpoints/adverse outcomes. However, this lack of information can be addressed by systematic documentation of causal relationships between measured biological effects and adverse outcomes, organised according to the AOP framework, and the complementary use of effect and exposure monitoring, as shown in the **Figure 1.1** below.

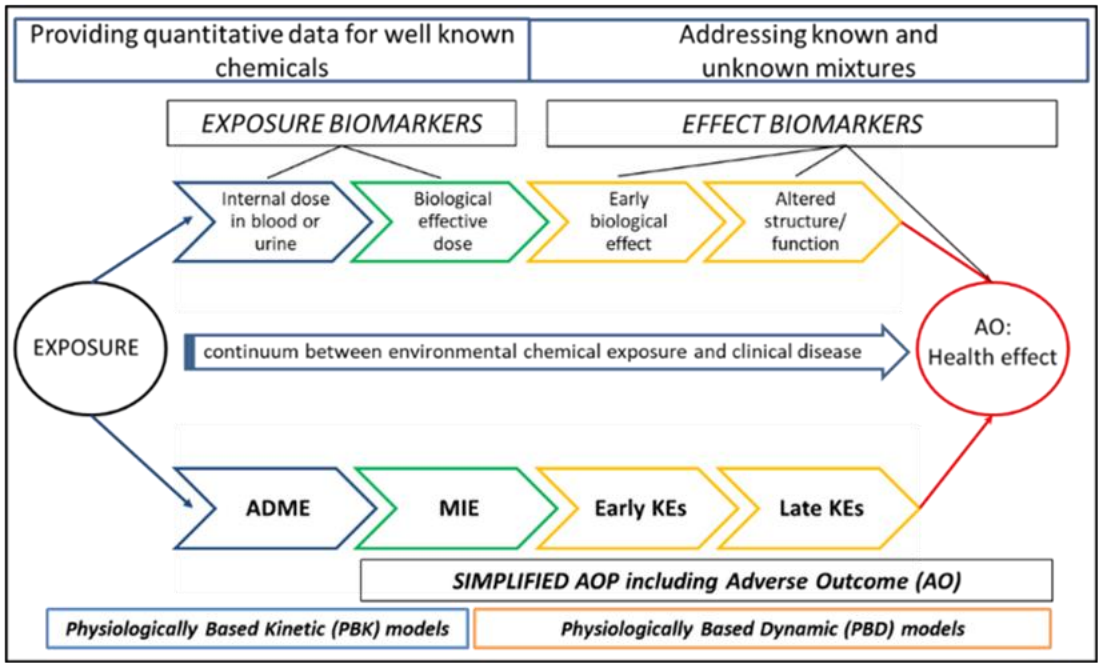


Figure 1 Integrative conceptual framework from exposure to adverse outcome, including the roles of biomarkers of exposure and effect, Adverse Outcome Pathways (AOP), and physiologically based kinetic and dynamic (PBK/D) models

ADME*: absorption, distribution, metabolism, excretion. MIE* : Molecular Initiating Event. KEs*: Key Events, AO : Adverse Outcome (adapted from Zare Jeddi et al. (2021)). MoA = ADME + AOP.

As a first step in developing these guiding principles, a prioritization survey requesting identification of effect biomarkers based on established MoAs was circulated to the OECD Working Party on Exposure Assessment (WPEA), OECD Working Party on Hazard Assessment (WPHA), and identified experts in the area of AOPs and effect biomarkers. Survey responses were received from 54 experts representing 35 international organisations. The survey results were discussed at an initial project kick-off meeting, from which potential effect biomarkers were identified within four main hazard classes: carcinogenicity (including genotoxicity and oxidative stress), estrogenic endocrine disruption, neurotoxicity & developmental neurotoxicity, and reproductive toxicity. For each of these hazard classes, expert groups were convened to evaluate the most relevant effect biomarkers, and to provide case studies in which mixture thresholds were derived by expert sub-groups (**see Annex A**).

A total of eleven effect biomarkers were characterized (**see Annex B**), but due to time and resource constraints, only seven biomarkers were assessed in case studies which also included derivation of mixture thresholds. The aim was to develop an approach for deriving thresholds based on a variety of MoAs that could be applied to other biomarkers.

Relevance of biomonitoring human effect biomarkers in risk assessment

Effect biomarkers quantify the effects of combined exposure to chemical mixtures from various sources and routes, in occupational settings and within the general population. They can thus identify subpopulations at higher risk and enable prioritization of new or improved risk management measures. Thus, integrating effect biomarkers into human biomonitoring (HBM*) programs in occupational and population safety assessments will support new or revised risk management measures. Where data are

available for both exposure and effect biomarkers in the general population, statistical analyses can be conducted to identify associations/correlations between these biomarkers. Potential covariates such as age, body mass index, sex, geographical locations and sociodemographic variables (including education, income and ethnicity) can be included in these analyses. This approach aids in interpreting trends in chemical exposure within the general population and helps identify vulnerable subpopulations (Borghese et al. 2022).

Vulnerable subpopulations are defined as groups of individuals within a population who, due to greater susceptibility or higher exposure, may be at an increased risk of experiencing adverse health effects from exposure to substances (HealthCanada 2022). To effectively leverage the use of effect biomarkers, it is imperative to assess their response level in quantitative risk assessment context. Occupational exposures differ from the general population with workers generally having exposure intensities greater than the general population, and shorter exposure duration (frequently assumed to be 8-hours per day versus 24 hour per day for the general population). In addition, workers are typically healthy adults, while the general population includes individuals who may be more sensitive to chemical exposures such as infants, toddlers, teenagers, aging populations, pregnant women and sick individuals. The guiding principles in this document are intended for occupational risk assessments, due to the higher expected exposures at workplaces compared to the general population. Robust risk assessments are essential for implementing effective occupational risk management measures (RMM*). For this reason, we have used terminology from the OECD Occupational Biomonitoring Guidance (OECD 2022). Our guidance is a practical approach for using effect biomarkers to derive mixture thresholds for occupational assessments and for their interpretation (see **Table 1**).

Use of effect biomarkers in an occupational setting should follow the same ethical standards as in exposure biomonitoring, which are described in detail in (Hopf et al. 2025; Hopf et al. 2024). In many national and international regulations, the employer has a responsibility to protect the health of their employees. Effect biomarkers can be used to identify elevated risks associated with hazardous working conditions. For example, on group level associated with incorrect use of Personal Protective Equipment (PPE*), behaviour at the workplace, or specific physiological status of workers. Reasons for elevated risks can be identified by a competent Occupational Health or Hygiene Professional (OHP*) and risk mitigation decisions can be taken to protect workers at risk. It is important to note that labelling individuals as vulnerable or marginalized may negatively impact their mental health and career options. Therefore, occupational biomonitoring must be based on sound scientific and ethical guidelines, allowing better protection for every worker. For this reason, this guidance describes an effect biomarker monitoring strategy using the similar exposure group (SEG*) approach. This is a comprehensive, worker-centric method for assessing and managing occupational exposures, potentially improving worker health and safety through a nuanced understanding of the biological impact of workplace hazards.

Relevance of biomonitoring environmental effect biomarkers in risk assessment

Complex mixtures of chemicals end up in environmental compartments (i.e. air, soil, water, biota) and may pose a risk to the environment as well as to human health. One of the most important compartments is the aquatic environment. Contamination of environmental waters cannot be addressed with chemical analysis alone (Dopp et al. 2019). The use of Effect-Based Methods (EBM*s) provides a straightforward approach to evaluate the risks from chemical mixtures in ecological systems, including mixtures of chemicals with the same MoA and complex mixtures. There are differences in sensitivity and exposure among different taxonomic groups based on both life history (which can impact exposure) and physiology (which can impact sensitivity). *In vitro* and/or *in vivo* effect biomarkers (EBM) can be useful to evaluate chemical mixture risk in a variety of taxa under different chemical pressures. EBMs could also be considered for the assessment of chemical status linked to a specific MoA (e.g. dioxin-like effects, estrogenicity) and/or can be included

as part of the determination of ecological status (EU Report 2021). Therefore, a common methodology for the derivation of mixture thresholds (also called Effect-Based Trigger (EBT*) values) is needed.

2 The derivation of mixture thresholds

Use of effect biomarkers in regulatory risk assessment requires establishing thresholds that encompass both ecological and human health risks. Thresholds values, known as Effect-Based Trigger values (EBT) or Occupational Biomonitoring Effect Levels (OBEL*). A common derivation method will allow inter-regulatory cost reductions on national and international level and can promote consistent standards across regulatory agencies.

Proposed principles for mixture threshold derivation for effect biomarkers

Four selection criteria were developed to assess environmental and human toxicological case studies using effect biomarkers, linking them to Adverse Outcomes (AO) and protection levels:

- Effect-biomarkers associated with Key Events (KE) or with adverse outcomes (AO) in one or more well supported AOPs*, e.g. with well understood toxicity pathway, that has undergone peer review or supporting weight of evidence relationships.
- Availability of reference compounds which act as prototypical stressors
- Opportunities to establish quantitative KE/AO relationships or risk classes for mixtures
- Datasets allowing defining thresholds or risk classes for human and ecological protection

When establishing mixture thresholds, it is important to apply principles similar to those used in characterizing effect biomarkers to ensure a high level of quality control and to enable the use of mixture thresholds in risk assessments. Such principles include:

Relevance & adversity: The effect biomarker must be relevant and plausible for the regulatory purpose. This can be achieved by a plausible link, as quantitative as possible, to a potential adverse effect.

Quantifiability: The effect-biomarker response can be measured either as a Bioanalytical Equivalent Concentration (BEQ*) for a MoA specific prototypical stressor or directly as an effect-level which can be linked to adverse effects or criteria or risk classes.

Robustness & repeatability: Effect biomarker measurements should demonstrate an acceptable level of intra- and inter-laboratory accuracy and precision, based on defined chemical reference standards.

Sensitivity & protectiveness: An effect biomarker response can be induced by a single substance or known and unknown mixtures (see **Figure 1.1**). The detection range of an effect-biomarker should be able to quantify response levels leading to adverse effects of single substances or mixtures in order to be sufficiently protective. For example, a Limit of Quantification (LOQ*) of an effect-biomarker should be below an accepted adversity level Environmental Quality Standard (EQS*) or Occupational Biomonitoring Level (OBL*) for a MoA specific stressor to ensure that measured mixture effects will not be underestimated.

Refinement & improvability: As more information is acquired, the probability that a measured effect biomarker will lead to an adverse outcome can be more precisely defined (same applies for exposure biomonitoring assessment values).

Interpretability & translatability: Different meaningful options in the derivation of effect biomarkers should be elaborated enabling targeted interpretation (see **Table 1 and 3)** which can be translated in appropriate risk assessment options and RMM.

The implementation of effect biomarkers needs guidance in the context of HBM programs or occupational studies. It is therefore necessary to pay particular attention to the detailed definition of their properties (Rodriguez et al. 2023). The same applies for Environmental Biomonitoring (EU Report 2021).

To address the lack of guidance for the use of effect biomarkers in HBM or occupational studies, a tiered approach for occupational biomonitoring was elaborated OECD Occupational Biomonitoring Guidance (OECD 2022). In this follow-up work, the initial OECD guidance has been further adapted to establish thresholds values needed for regulatory risk assessment, e.g. Occupational Biomonitoring Effect Levels (OBEL) for humans and Effect-Based Trigger values (EBT) for ecological risk assessment. For occupational (exposure) biomonitoring, threshold derivation has been described in the OECD Occupational Biomonitoring Guidance (Hopf et al. 2025; Hopf et al. 2024; OECD 2022). We used the same principles for effect biomarkers as for exposure biomonitoring, in order to promote a harmonized approach, save resources, and enhance consistency across regulatory agencies. The derivations of threshold levels are described in the individual selected case studies (**Annexes A1-A7**).

Depending on availability of effect-biomarker and toxicological data for their assessment, four tiers (0 – 3) are defined and can be assessed as threshold levels (technical, reference, provisional, and refined) (**Table 1 and 3)**. The Refined Level (Tier 3) provides the highest level of certainty about health risk. The meaning of threshold exceedance differs by level and will guide appropriate risk assessment options and RMM, which are not specified in this document.

An OBEL or EBT can be linked directly to an adverse effect level or a risk of adverse outcome, exceedance indicates a need for risk management action. A POBEL or P-EBT would correspond to a MoA -specific accepted Exposure Level (OBL), EQS*, or risk class indicates to an increased probability of adverse effects. At this level a mixture-related health risk cannot be excluded (see **Figure 1.1** & **Table 1**). A ROBEL or R-EBT indicates the 5%/Median/95% reference level of an effect biomarker. A TOBEL and T-EBT provides a technical assessment level with physiological relevance (see explanation below **Table 2**). An interpretation of assessment levels is provided in **Table 3** for detection of effects, their quantification and interpretation for MoA related risks.

Table 1. Proposed concept of Occupational Biomonitoring Effect Level (OBEL) and for Effect Based Trigger values (EBT) allowing their interpretation.

Human Health (HH)	Environmental Health (ENV)	Tier	Level	Meaning if exceeded
OBEL=Occupational Biomonitoring Effect Level	**EBT =Effect-Based Trigger value**	**3**	**Refined**	**Health risk indicated**
POBEL	**P-EBT**	**2**	**Provisional**	**Health risk may be indicated**
ROBEL	**R-EBT**	**1**	**Reference**	**Exposure above or below a reference level (e.g.>95%, <5%)**
TOBEL	**T-EBT**	**0**	**Technical**	**Exposure indicated or not (e.g. >LOQ* or >LOD*)**

Similar principles to those established for exposure biomonitoring quality assessment should be applied to effect biomonitoring used for risk assessment. The majority of the effect biomarkers presented in the case studies (**Annexes A1-A7**) can be applied in an occupational context (see related guiding principles document "Guiding principles to advance occupational mixture risk assessment with effect biomarkers." (OECD 2025).

Guiding principles are summarized briefly in the following paragraph:

As mentioned, traditional occupational health risk assessments have relied on external exposure measurements, such as air monitoring, which may not fully capture the complexities of workplace exposures. Building on established approaches used for many decades in air monitoring, the related guiding principles document advocates for the integration of effect biomarkers and the use of Similar Exposure Groups (SEGs*) to enhance the accuracy and efficiency of occupational exposure assessments. Effect biomarkers accounts for individual susceptibility, all exposure routes, cumulative effects over time, and the impact of chemical mixtures. The SEG approach refines risk assessments by focusing on group-level exposures, enhances data interpretation through statistical robustness, and facilitates targeted interventions for groups rather than individuals. Collecting appropriate exposure data is crucial for effect biomonitoring including defining SEGs. SEGs are created based on evaluating work processes and exposure determinants - factors that influence the levels (intensity), frequency, and duration of exposure to chemicals in the workplace, aiming-to create homogeneous groups regarding exposure potential; and developing a sampling strategy. When using effect biomonitoring, additional information on individual worker characteristics such as sociodemographics and health status should also be collected to adjust for factors that influence susceptibility. There is the need to match the exposed study group with a control group and to inform about possible exclusion criteria. A tiered system with four assessment levels supports interpreting effect biomarker data (**Table 1+3**). This framework translates raw biomarker data into actionable insights for managing workplace risks. By conducting effect biomarkers monitoring with the SEG approach, this strategy offers a more comprehensive, worker-centric method for assessing and managing occupational exposures, potentially improving worker health and safety through a nuanced understanding

of the biological impact of workplace hazards. Although comparisons are performed at the group level, single measurements should be not ignored. Individual factors may affect the results (e.g. pre-existing diseases, medications, alcohol consumption), this should be addressed by questionnaires for all involved workers and in the control group.

3 Case studies demonstrating the derivation of Occupational Biomonitoring Effect Levels (OBEL) and Effect-Based Trigger values (EBT)

Seven case studies were developed, following the project priorities and covering ecological (estrogenic endocrine disruption) and human health (carcinogenicity, including genotoxicity and oxidative stress, neurotoxicity & developmental neurotoxicity, and reproductive toxicity). The case studies are briefly summarized here and are described in full **in Annexes A1-A7**.

Summary A1-A2) Estrogenic Endocrine Disruption in water assessment (ENV): Vitellogenin induction in male fish & Estrogen Receptor alpha transactivitation

Two effect biomarkers, estrogen receptor alpha (ERα) transactivation (measured *in vitro*) and vitellogenin induction in male fish (measured *in vivo*), were evaluated in this case study to assess aquatic exposure to exogenous estrogenic chemicals (i.e., chemicals with estrogen-like hormonal properties).

Since the alteration of certain hormonal processes in humans and wildlife was first reported and linked to long-term exposures to low concentrations of estrogenic chemicals and their mixtures, erogenicity is one of the most well-studied endpoints related to endocrine disruption. ERα transactivation is a key event in estrogenic AOPs and is associated with reproductive or developmental effects. The occurrence of significant vitellogenin induction in male fish may signal a high probability of adverse estrogenic hazard. Using the template for characterization of estrogenic effect biomarkers, both endpoints and test methods showed favourable characteristics for their use as effect biomarkers in environmental water quality monitoring.

EBTs corresponding to technical limits of detection (technical-EBT), significant elevation of response above reference or baseline conditions (reference-EBT), potential environmental health risk (provisional-EBT), or high probability of health risk (refined-EBT) were proposed for both ERα-transactivation and plasma vitellogenin induction in male fish. Using this approach, EBTs ranging from 0.18-0.56 ng/L 17β-estradiol equivalents were derived for five separate ERα transactivation assays. Likewise, EBTs ranging from 60-3000 µg vitellogenin/mL plasma were derived for six different fish species (*Pimephales promelas, Danio rerio, Oryzias latipes, Oncorhynchus mykiss Cyprinus carpio, Micropterus salmoides*). Potential use of these proposed EBTs is illustrated via application to multiple case studies including laboratory-based evaluation of a single chemical exposure, *in vitro* and *in situ* exposure to treated wastewater, and *in vitro* and in situ exposures to ambient surface water.

The results contribute to ongoing international efforts to support risk assessment of complex mixtures using effect biomarkers anchored to adverse outcome pathways (AOP). The defined thresholds fall within a fairly narrow range across six fish species and five transactivation assays, indicating good transferability between species and *in vitro* methods, supporting read-across opportunities and grouping similar modes of action and biological endpoint characteristics.

Summary A3) Genotoxicity mixture threshold assessment in Occupational Biomonitoring via the lymphocyte Cytokinesis-block micronucleus (L-CBMN*) assay

There are several assays available to detect genotoxicity in human studies. We selected the micronucleus (MN*) assay because of the association between the outcome of the assay and increased risk of cancer, supported by AOPs and prospective cohort studies that have found significant associations between overall cancer incidence and an increased MN frequency in healthy subjects. We derived human health risk levels for MN based on the L-CBMN assay. More specifically, for the ROBEL evaluation, we propose to establish a baseline MN frequency specific for each laboratory, taking into account major confounders (i.e., age, sex and interlaboratory variability) among non-exposed subjects. Additionally, it is recommended to minimise, interlaboratory variability e.g. by using harmonized protocols for sample processing and analysis and regular interlaboratory comparisons. For the sake of feasibility (quite often the database of one laboratory is small), the age classes for establishing baseline MN frequency can be large (e.g., ≤50 years; > 50 years). For the OBEL - to identify those subjects for which a health risk is indicated – we propose to assess the micronuclei mean ratio (MN-MR) comparing MN mean frequency in the exposed group with respect to a control group (or historical baseline) frequency, adjusted by sex and by age-class. According to evidence from various large studies and meta-analyses we advise that whenever there is a statistically significant increase of two-fold (MN-MR≥2.0) in exposed workers vs matched unexposed controls, detailed exposure assessment to identify potential causative agents and control measures to mitigate the risk need to be undertaken. We acknowledge the inherent limitations of the L-CBMN assay, and we see promise for other assays to detect MN frequencies, such as in exfoliated buccal cells or adding improvement of analytical methods in L-CBMN e.g. flow cytometry, Artificial Intelligence image analysis. However, until other assays are associated to cancer to the same extent as the L-CBMN assay, these other genotoxicity assays can only be used supportive for the risk assessment approach as outlined in this work. Moreover, in this case study for L-CBMN (see **Annex A3)** practical examples are given how to apply the L-CBMN assay, and which additional confounders need to be addressed.

Summaries A4-A6) (Developmental)-Neurotoxicity assessments in occupational context

Summary A4) Brain Derived Neurotrophic Factor (BDNF*) assessment in serum

Brain derived neurotrophic factor (BDNF) plays a key role in neuronal survival, differentiation, maturation and brain plasticity. Peripheral BDNF gene and protein levels are used as potential biomarkers for learning and memory deficits in children and adults as well as cognitive performance in depression, Parkinson and Alzheimer's disease. Based on epidemiological and clinical studies, BDNF DNA methylation is also proposed as reliable biomarker of neurological effect, being more stable over time compared to gene expression or protein levels. TOBEL: Total BDNF protein levels are frequently and preferably measured in serum using enzyme-linked immunosorbent assay (ELISA) that quantify levels in nearly 100% of human samples with an analytical LOQ between 0.03 and 62.5 pg/mL, depending on the specific assay. ROBEL:

Reliable serum BDNF levels, measured in populations are available in epidemiological studies. In adults, mean total serum BDNF ranges from 32.3 to 40 µg/L and in adolescents, mean total serum BDNF ranges from 26.8 to 31.5 µg/L. As a ROBEL, a serum BDNF level of 33 µg/L can be used. POBEL: Lead (Pb) is the reference chemical since many animals and some epidemiological studies have linked Pb exposure to alterations in BDNF levels and neurotoxicity. In animals, the MoA for altered BDNF levels in association with Pb exposure is well documented in the AOP -Wiki. Antagonism of Pb on N-methyl-D-Aspartate receptors is a molecular initiating event leading to altered BDNF regulation in the brain that may result in learning and memory impairments, as described in AOPs 12 and 13. Occupational exposure to manganese showed a dose-response inverse relationship with plasma BDNF levels resulting in cognitive deficits. Currently, the German biological tolerance value for occupational exposures (so called BAT value) defined as 150 µg Pb/L in blood is not protective for the offspring of female lead-exposed workers. Bringing all lines of evidence together, a 30-50% relative reduction in BDNF levels is associated with impaired learning and memory function in worker and animal studies. Severe neurotoxic effects are reported by less than 50% BDNF reduction. As intermediate value a BDNF level of 16.5 µg/L in serum can be used. Comparisons should be mainly performed on group level, single measurements should be not ignored. OBEL: Many epidemiological studies have confirmed the association between blood Pb levels and decreased neuronal function or cognitive impairment. An occupational study allowing a correlation between Pb exposure and cognitive impairment to derive a refined OBEL is lacking.

Summary A5) Neurofilament light chain (NfL*) measured in serum

Neurofilaments are structural scaffolding proteins that are exclusively expressed in neurons, mainly in axons, but expression in other compartments of neurons (e.g. synapses) has also been reported. A variety of neurotoxic compounds, stroke, traumatic brain injury, neurodegenerative and psychiatric disease are associated with neuroaxonal damages reducing the microstructural integrity of the nervous system. Such microstructural damages impairing the overall functioning of the brain can be detected by elevated neurofilament light chain (NfL) levels in blood. Neurotoxic side-effects of anti-cancer drugs are known to increase serum levels of NfL. Neurofilaments are not explicitly mentioned in AOPs but several key events, such as “synaptogenesis, decreased” or “neuronal network function, decreased” are directly related to the integrity and dynamics of such scaffolding proteins in axons, dendrites and synapses. TOBEL: NfL is usually measured by commercial ELISA kits and the single-molecule array (SiMoA) assay has the lowest LOQ. ROBEL: NfL is routinely used in various clinical applications and large database including healthy controls are available. A general cutoff level is 10 ng/L (equivalent to pg/mL) in healthy subjects. POBEL: Based on age-adjusted population data, the 95%-percentiles of unexposed subjects can be used to identify workers that are “at risk” of developing cognitive impairments and their working environment should be checked for possible neurotoxic exposures. OBEL: Reliable and valid data relating biomarkers of exposure to known neurotoxicants, such as blood Pb levels to NfL levels are currently not available. If such data become available a well-established NOAEL, e.g. the BEI of 150 µg Pb/L, can be related to NfL levels in exposed workers.

Summary A6) Neurogranin (NGRN, Ng*) assessment in serum

Neurogranin (Ng) is a postsynaptic protein in excitatory neurons essential for synaptic plasticity and memory. Upon NMDA receptor activation, it releases calmodulin, triggering calcium-dependent pathways. Reduced Ng in the hippocampus and frontal cortex of Alzheimer’s disease (AD) patients indicates early synaptic loss and cognitive decline. Elevated cerebrospinal fluid (CSF) Ng levels correlate with hippocampal atrophy in AD and mild cognitive impairment (MCI) patients. In cases of traumatic brain injury, stroke, brain infection, epilepsy, and carbon monoxide (CO) poisoning, Ng also reflects synaptic degeneration and is a potential biomarker for synaptic dysfunction and cognitive impairment. Developmental exposure to chemicals can reduce Ng expression in the rat hippocampus. TOBEL: Measurement of Ng levels typically involve ELISA and electrochemiluminescence techniques in CSF and

plasma/serum. ROBEL: Ng levels in CSF increase with age and are higher in females. In adults, serum Ng levels average 14.9±30.3 ng/mL, but robust reference intervals and cut-off values are lacking. POBEL: Occupational and environmental exposure to neurotoxic chemicals may cause neurodegenerative diseases. Changes in Ng levels can indicate synaptic degeneration and serve as biomarker for early prediction of cognitive impairment in chemically induced neurodegeneration. Animal studies showed that metal exposure affects Ng expression. Current German biological tolerance value for occupational exposures (BAT value) of 150 µg Pb/L in blood is not protective for offspring of female lead-exposed workers. OBEL: Despite not being a KE in the AOPs, Ng is mentioned in AOP 42 and KE 756, supporting "altered hippocampal gene expression". While ELISA is the most common method for measuring Ng levels, many studies lack detailed information on its analytical performance, often focusing solely on diagnostic performance. Validation criteria include antibody details, LOD, LOQ coefficients of variation, recovery, and calibration curves. Only two studies met most criteria, limiting validity. Variability of control groups across studies by age, sex, race, country and other factors hinder comparability and reference value identification.

Summary A7) Reproduction toxicity induced via low testosterone levels in male humans

Testosterone, the primary male sex steroid hormone, is crucial for sex differentiation, male characteristics, spermatogenesis, and fertility. It is mostly bound to proteins in the blood, such as sex hormone-binding globulin (SHBG*) and albumin, with only a small fraction freely available. Both albumin-bound and free testosterone are considered biologically active testosterone. Total (SHBG- and albumin-bound and unbound) serum testosterone (TT*) levels can indicate reproductive health, with low levels linked to adverse outcomes such as decreased fertility and abnormal sexual differentiation during development. The use of TT levels as an effect biomarker for assessing reproductive toxicity in occupational biomonitoring is supported by various studies. The biological relevance of TT levels is based on its role in several AOPs related to male reproductive health. These AOPs provide a framework for understanding how environmental and occupational exposures can disrupt endocrine function, leading to reduced TT levels and potential reproductive harm. ROBEL: TT level is age dependent and well investigated. Reliable serum reference levels are available in the general population (**Table 2**). Significant decreases below reference levels may indicate exposure to TT reducing substances. Human TT level is a sensitive and well investigated clinical biomarker of effect. This effect biomarker can be used to identify in collective assessments, occupational exposures that reduce testosterone levels. We assume there is a correlation between TT levels and adverse sperm quality parameters as defined by World Health Organization (WHO), thus TT levels could be used to develop POBEL, although access to comprehensive datasets, such as those from the WHO, is needed for further analysis.

4 Compilation of mixture thresholds and their interpretation

Based on the case studies conducted by the expert groups, followed by multiple rounds of comments discussions and review meetings, mixture threshold levels for the seven effect-biomarkers were derived within the expert groups. As in other assessments the confidence (reliability, relevance and weight of evidence) of toxicological data was intensively assessed. All key studies and datasets were discussed with an expert group. Overall, four expert drafting groups were involved in the derivation and in several rounds of comments. The derivation outcomes were discussed in the large project meetings, providing additional safety. For L-CBMN large epidemiological datasets were available with correlations to several cancer and non-cancer diseases. If the data base was not strong enough to derive refined OBEL then only POBEL or ROBEL were derived. This was the case for BDNF, NfL with sufficient confidence to derive a POBEL. For TT and Ng the database had only sufficient evidence to derive reference levels. The link to adverse effects of these effect biomarkers is likely, but is currently missing in a quantitative way. These results are summarized in **Table 2**; and details are given in **Annexes (A1-A7**).

Table 2. Mixture thresholds assessment levels and functions for seven effect-biomarkers for occupational and environmental assessments.

Abbreviations for terminology: Occupational Biomonitoring Effect Level (OBEL); Effect Based Trigger values (EBT) (specific abbreviations see Table 1 and Glossary)

Abbreviations for effect biomarkers: Vitellogenin (VTG*), Estrogen Receptor alpha Transactivation Assay (ERα TA); Brain Derived Neurotrophic Factor (BDNF); Neurofilament Light Chain (NfL), Neurogranin (Ng), Lymphocytes Cytokinesis Block Assay(L-CBMN), Testosterone Level (TT)

Effect-biomarker abbreviation	Assessment level and function			
	TOBEL or T-EBT **-Technical limit to quantify exposure and effect**	**ROBEL or R-EBT** **-Reference to interpret exposure and effect**	**POBEL or P-EBT** **-Provisional to quantify likely health risks**	**OBEL or EBT** **-Refined to quantify direct health risks**
VTG[#]	2 µg/mL	12 µg/mL	≈120 µg/mL	≈1200 µg/mL
ERα TA[##]	LOD: <0.01 ng E2-EQ/L[###] LOQ: 0.01 - 0.1 ng E2-EQ/L	surface water dependent	0.40 ng E2-EQ/L	0.28 ng E2-EQ/L
BDNF	The LOQ for BDNF is between 7.8 and 62.5 ng/L, with Simoa technology a LOQ of 0.03 ng/L is reached	In adults, mean BDNF levels of 33 ± 8.5-17 µg/L	Severe neurotoxic effects are reported by less than 50% BDNF reduction in worker and animal studies. As intermediate value a BDNF level of 16.5 µg/L in serum can be used.	not assessed
NfL	LOQ of 0.02 ng/L NfL	Extensive clinical datasets are already available providing normative data. A fixed cutoff level would be >10 ng/L measured in serum	NfL levels above age specific can lead to reduced structural integrity of the brain: age classes NfL [years]→ (95% percentile [ng/L])	not assessed

			41-50--> 12-15 51-60--> 15-20 61-70--> 20-26 71-80--> 26-32	
Ng	LOQ of 0.05 ng/L Ng (ELISA commercial kit)	In adults (non-diseased controls), serum Ng is usually below 200 µg/L	not assessed	not assessed
L-CBMN	MN[####] ratio can always be quantified, no limitation is given. LOQ is only cell count dependent.	Adjust the control group for age and sex (for example large age-classes can be used, e.g., ≤50 years; > 50 years)	not needed	equal or greater than 2 fold-increase in comparison to control group may indicate a cancer and/or non-cancer disease risk
TT -Level	LOQ is around 20-50 ng/L in blood	age classes TT [years]→ (5% percentile [ng/L] for n = 9054) 19-39 →2730 40-49 →2430 50-59 →2220 60-69 →2210	not assessed	not assessed

[#]Male plasma vitellogenin (VTG) induction assessed for male *Pimephales promelas, also assessed for Danio rerio, Oryzias latipes, Oncorhynchus mykiss, Cyprinus carpio Micropterus salmoides (*see **Annex A2**)

[##]*Estrogen Receptor alpha Transactivation Assay (ERα TA) assessed for ERα-CALUX, also assessed for MELN, ERα-GeneBLAzer, HeLa-9903, p-YES (*see **Annex A1***)*

[###]E2-EQ/L =: estradiol equivalents per liter

[####]micronuclei (MN)

Please note: A TOBEL or T-EBT indices the technical limit to quantify an exposure effect but needs to be checked for sensitivity in advance with the other assessment levels. In contrast to a TOBL for an exposure biomarker limiting exposure to a technically achievable level, a TOBEL should always be more sensitive than a ROBEL, POBEL or OBEL to ensure that the response can be interpreted in a physiologically relevant range and before health risks are indicated.

Table 3. General traffic light system for effect biomarker interpretation

Abbreviations for terminology: Occupational Biomonitoring Effect Level (OBEL); Effect Based Trigger values (EBT) (specific abbreviations see Table 1); Mode of action (MoA)

Detection result (response)	**Quantification[#]**	**Interpretation for MoA related risk**
No measurable exposure and effect	Response < or > TOBEL or T-EBT	Risks for adverse outcome effects are unlikely if effect biomarker is sensitive and MoA specific
Exposure and effects deviate from reference level	Response > or < ROBEL or R-EBT	Attention: Health risks cannot be excluded
Exposure and effects differ from provisional health risk levels	Response > or < POBEL or P-EBT	Warning: Risks for adverse outcome effects are likely
Exposure and effects differ from refined health risk levels	Response > or < OBEL or EBT	Severe warning: Measured effects-levels are known to lead to adverse health outcomes

[#]For occupational effect biomarker assessments, additional guiding principles are available, containing minimum statistical requirements and reporting recommendations; see the OECD “Guiding principles to advance occupational mixture risk assessment with effect biomarkers.(0ECD 2025)”

5 Lessons learned from case study assessments & outlook

All effect-biomarkers assessed in the case studies can identify complex exposures with sufficient sensitivity and enable to interpret MoA specific effects in a physiological relevant range. This was needed to fulfill the **sensitivity** and **protectiveness** criteria. Therefore, these effect-biomarker can quantify known and unknown (not measured and not measurable) exposure effects meaningful for risk assessment. All responses are **quantifiable** enabling the monitoring of exposures and their related effects over time. Using effect-biomarkers can lower the risk of having health-relevant effects from not analysed substances. For five of seven effect-biomarkers it is also possible to **quantify** likely and direct health effects in an AOP compatible concept. This ensures that there is plausible **relevance** in terms of **adversity.**

In summary our tiered assessment system is practical, robust and **interpretable & translatable** for all effect biomarkers (see **Table 2 & 3**). Additionally, the proposed assessment concept is **transferrable** to many other effect biomarkers, for example for estrogenic assessments this evaluation progress can be multiplied by a factor of five to six for other species and assays (**see Annex A1-2**).

In occupational context correlation studies with BEQ and adverse health endpoints or OBL are needed (examples, Neurotoxicity and developmental Neurotoxicity). Such studies would strengthen the link between exposure and effect biomonitoring more effectively, enabling for **refinement & improvability** in risk assessment. Interlab-validation is in some cases needed to have more comparability of results for example genotoxicity and oxidative stress. In contrast to this for the effect-biomarkers NfL and TT -level sufficient assessment knowledge is already available to allow an age specific assessment.

Additionally in parallel developed, guiding principles for occupational mixture risk assessments (see OECD 2025 “Guiding principles to advance occupational mixture risk assessment with effect biomarkers”) enable collective assessment levels to cope with interindividual variabilities of effect biomonitoring providing, improving **robustness and ensuring repeatability**. We also recognized that a risk class system for POBL and P-EBT can be a good assessment option to cope with missing correlation studies and the remaining uncertainties. The proposed assessment concept allows generally a **systematic interpretation** of effect-biomonitoring responses in all cases on different assessment levels. Assessments on higher tiers will become more applicable with increasing AOP knowledge and quantifiable connections to adverse endpoints (example human male TT level). Overall, with the different selected case studies for quite different MoAs, we have identified an enormous potential and opportunities to establish quantitative KE/AO relationships or risk classes for mixtures, paving the way for the use of additional effect-biomarkers in the future. It is likely that the growing AOP knowledge will enable quantitative assessments and more mixture assessments in the future.

6 Overall conclusions & recommendations

Conclusions

As mentioned in the beginning presently, effect biomarkers are the most direct option to address the risk of known and unknown mixtures in an integrative way. Therefore, we concluded that the use of effect biomarkers should be enabled and encouraged together with a systematic assessment concept. Our proposed assessment concept for effect biomarker works for mixture assessments and is AOP and risk assessment compatible (see **Tables 1-3** and **Annex A**). It allows users to quantify the effects of exposure to known and unknown mixtures in an evidence-based health risk assessment for a variety of effect biomarkers. Depending on the available knowledge, all effect biomarkers can in principle be systematically assessed and interpreted. With increasing knowledge, this approach enables a more effective bridging of effect and exposure biomarker assessments, allowing for a more quantitative evaluation of adverse health risks. In the project related case studies, we have assessed various effect-biomarkers and provided mixture threshold derivations. Generally, our assessment system has shown applicability for environmental and human health related effect biomarkers in risk assessment following the one health concept.

Recommendations

It is recommended that mixture thresholds should be derived for more effect biomarkers to enable broader interventions focused on different mixture risks. The necessary knowledge should be generated in a targeted manner for relevant mixtures posing adverse health risks to humans or the environment. Additionally, the identification of new effect-biomarkers is highly recommended, and our document provides a clear strategy to check their applicability. It is necessary to characterize effect biomarkers in a systematic way (**see Annex B**) to identify their potential for risk assessment and to expand AOP knowledge to both establish the causal relationship between effect biomarkers and potential for adverse outcomes and, where possible, to better define the quantitative thresholds of response that signal adversity. For some effect biomarkers, further interlaboratory -validation may be needed, to allow for a robust comparison of effect biomarker data across studies. Generally, we recommend a systematic application of effect biomarkers in situations in which potentially hazardous and/or complex exposures are likely (see relevant triggers for effect-biomarker use in occupational assessments in **Annex C**). This should preferably be done in combination with exposure biomarker biomonitoring campaigns to improve their interpretation and accuracies. This will increase understanding of the proportion of mixture effects which can be explained by selected exposure biomonitoring analytes. The general assessment concept of mixture thresholds is recommended for further application in effect-biomonitoring guidance (e.g. (EFSA et al. 2024)) to enable an integrative risk assessment solution.

7 References

Beronius, A., P. Nymark and E. Wincent (2021), *Adverse Outcome Pathways – principles and applications.*, Karolinska Institute of Environmental Medicine., https://ki.se/media/255075/download.

Boran, T. et al. (2024), "The cyclin-dependent kinase inhibitor abemaciclib-induced hepatotoxicity: Insight on the molecular mechanisms in HepG2/THP-1 co-culture model", *Toxicology Letters*, Vol. 391, pp. 1-12, https://doi.org/10.1016/j.toxlet.2023.11.005.

Borghese, M. et al. (2022), "Individual and mixture associations of perfluoroalkyl substances on liver function biomarkers in the Canadian Health Measures Survey", *Environmental Health*, Vol. 21/1, https://doi.org/10.1186/s12940-022-00892-6.

Carmichael, N. et al. (2011), "Using mode of action information to improve regulatory decision-making: An ECETOC/ILSI RF/HESI workshop overview", *Critical Reviews in Toxicology*, Vol. 41/3, pp. 175-186, https://doi.org/10.3109/10408444.2010.541225.

Dopp, E. et al. (2019), "Effect-based methods in combination with state-of-the-art chemical analysis for assessment of water quality as integrated approach", *International Journal of Hygiene and Environmental Health*, Vol. 222/4, pp. 607-614, https://doi.org/10.1016/j.ijheh.2019.03.001.

Edwards, S. et al. (2016), "Adverse Outcome Pathways—Organizing Toxicological Information to Improve Decision Making", *The Journal of Pharmacology and Experimental Therapeutics*, Vol. 356/1, pp. 170-181, https://doi.org/10.1124/jpet.115.228239.

EUReport (2021), *Technical proposal for effect-based monitoring and assessment under the Water Framework Directive.*, European Commission and CIS partners, https://circabc.europa.eu/sd/a/5b2f9e69-e078-429f-94b7-b361d123e072/211013_EBM%20report_FINAL_WG_Chem_Oct_2021.pdf.

Health Canada (2022), *Consideration of vulnerable populations in risk assessment*, Government of Canada, https://www.canada.ca/en/health-canada/services/chemical-substances/fact-sheets/consideration-vulnerable-populations-risk-assessment.html.

Hopf, N. et al. (2025), "Introducing the OECD guidance document on occupational biomonitoring: A harmonized methodology for deriving occupational biomonitoring levels (OBL)", *Toxicology Letters*, Vol. 403, pp. 132-143, https://doi.org/10.1016/j.toxlet.2024.12.006.

Hopf, N. et al. (2024), "A harmonized occupational biomonitoring approach", *Environment International*, Vol. 191, p. 108990, https://doi.org/10.1016/j.envint.2024.108990.

OECD (2025), *Guiding principles to advance occupational ixture risk assessment with effect biomarkers*, OECD Series on Testing and Assessment, No. ***, OECD Publishing, Paris.

OECD (2022), *Occupational Biomonitoring Guidance Document*, OECD Series on Testing and Assessment, No. 370, OECD Publishing, Paris, https://doi.org/10.1787/11bc2c7a-en.

Rodríguez-Carrillo, A. et al. (2023), "Implementation of effect biomarkers in human biomonitoring studies: A systematic approach synergizing toxicological and epidemiological knowledge", *International Journal of Hygiene and Environmental Health*, Vol. 249, p. 114140, https://doi.org/10.1016/j.ijheh.2023.114140.

World Health Organization (WHO) (1993), *Biomarkers and risk assessment : concepts and principles*, the joint sponsorship of the United Nations environment Programme, the International Labour Organisation, and the World Health Organization, https://www.who.int/publications/i/item/9241571551, https://iris.who.int/items/f16ba693-72a8-46ae-8529-0d1f96743e4a/full.

Zare Jeddi, M. et al. (2021), "Towards a systematic use of effect biomarkers in population and occupational biomonitoring", *Environment International*, Vol. 146, p. 106257, https://doi.org/10.1016/j.envint.2020.106257.

8 Project participation

We would like to acknowledge the numerous colleagues and experts involved in project initiation, for their active involvement in case studies, for their active participation in project meetings and for their review work of project results. Especially, we would like to thank the hosts of project meeting events for their enormous hospitality. Project participants are listed in alphabetical order by last names in the following table.

Last name:	First name:	Organisation
Aimonen	Kukka	Finnish Institute of Occupational Health (FIOH), Finland
Angeli	Karine	French National Agency for Food, Environmental and Occupational Health & Safety (ANSES), France
Aoyagi	Tomoko	OECD
Baucher	Marie-Ange	OECD
Beronius	Anna	Karolinska Institute, Sweden
Bessems	Jos	Flemish Institute for Technological Research (VITO), Belgium
Bil	Wieneke	National Institute of Public Health and the Environment (RIVM), Netherlands
Bláha	Ludek	Research Centre for Toxic Compounds in the Environment (RECETOX), Czechia
Bloch	Denise	Federal Institute for Risk Assessment (BfR), Germany
Bonassi	Stefano	San Raffaele S.p.A., Italy
Bonefeld-Jørgensen	Eva Cecilie	University of Aarhus, Denmark
Bonvallot	Nathalie	Ecole des hautes études en santé publique (EHESP), France
Browne	Patience	OECD
Busch	Wibke	Helmholtz Centre for Environmental Research (UFZ), Germany
Chorley	Brian	U.S. Environmental Protection Agency (EPA)
Crepet	Amélie	Agence nationale de sécurité sanitaire (Anses), France
Delrue	Nathalie	OECD
Dopp	Elke	University Duisburg-Essen, Medical Faculty, Germany
Duca	Radu	National Health Laboratory, Luxembourg
Dusinska	Maria	Norwegian Institute for Air Research
Engelbrecht	Vera	PETA Science Consortium International e.V.
Escher	Beate	Helmholtz Centre for Environmental Research - UFZ, Germany
Farcal	Lucian	EFSA (European Food Safety Authority)
Fenech	Michael	University of South Australia
Fernandez	Marieta	University of Granada, Spain
FitzGerald	Rex	Universities of Basel and Geneva, Switzerland
Frey	Sabine	Federal Office of Public Health (FOPH), Switzerland
Gate	Laurent	Institut National de la Recherche et de la Sécurité (INRS), France
Ghosh	Manosij	KU Leuven, Belgium
Göen	Thomas	University of Erlangen-Nuremberg, Germany
Groh	Ksenia	Eawag (Swiss Federal Institute for Aquatic Science and Technology), Switzerland

Gutleb	Arno	Luxembourg Institute of Science & Technology (LIST), Luxembourg
Guyton	Kate Z.	US National Academies of Sciences, Engineering, and Medicine
Hassold	Enken	German Environment Agency (Umweltbundesamt – UBA), Germany
Hernandez-Jerez	Antonio	University of Granada, Spain
Hopf	Nancy	Center for Primary Care and Public Health (Unisanté), Switzerland
Hotta	Asako	OECD
Jagla	Magdalena	Environment and Climate Change Canada
Jones	Kate	Health and Safety Executive (HSE), UK
Kasiotis	Konstantinos	Benaki Phytopathological Institute (BPI), Greece
Kim	Yong Mu	Ministry of Food and Drug Safety,South Korea
Koller	Michael	Swiss National Accident Insurance Fund (SUVA), Switzerland
Kumar	Anupama	Commonwealth Scientific and Industrial Research Organisation (CSIRO)
Langer	Miriam	FHNW, Switzerland
Lecloux	Hélène	Federal Public Service (FPS) Health, Food Chain Safety and Environment, Belgium
Leusch	Frederic	Griffith University, Australia
Long	Manhai	University of Aarhus, Denmark
Louro	Henriqueta	National Institute of Health Dr. Ricardo Jorge (INSA), Portugal
Luijten	Mirjam	National Institute of Public Health and the Environment (RIVM), Netherlands
Machera	Kyriaki	Benaki Phytopathological Institute (BPI), Greece
Mizutani	Yoshitaka	Ministry of the Environment, Japan
Moeller	Ruth	Laboratoire nationale de santé, Luxembourg
Moritani	Naoko	OECD
Murata	Takaaki	Environmental Health Department, Ministry of the Environment, Japan
Mustieles	Vicente	University of Granada, Spain
Ndaw	Sophie	Institut National de la Recherche et de la Securite (INRS), France
Paini	Alicia	EFSA (European Food Safety Authority)
Palmen	Nicole	National Institute for Public Health and the Environment (RIVM), Netherlands
Palmont	Philippe	Agence nationale de sécurité sanitaire (Anses), France
Pasanen-Kase	Robert	State Secretariat for Economic Affairs (SECO), Switzerland
Pieper	Christina	Federal Institute for Risk Assessment (BfR), Germany
Plichta	Veronika	Austrian Agency of Health and Food Safety (AGES); Austria
Poddalgoda	Devika	Health Canada, Canada
Price	Anna	European Commission/DG Joint Research Centre
Probst-Hensch	Nicole	Swiss Tropical and Public Health Institute, Switzerland
Remy	Sylvie	Flemish Institute for Technological Research (VITO), Belgium
Rissler	Jörg	Institute for Occupational Safety and Health of the German Social Accident Insurance, Germany
Rodriguez Carrillo	Andrea	Flemish Institute for Technological Research (VITO), Belgium
Sams	Craig	Health and Safety Executive (HSE), United Kingdom
Santonen	Tiina	Finnish Institute of Occupational Health (FIOH), Finland
Schmeisser	Sebastian	German Federal Institute for Risk Assessment (BfR)
Schmitz-Spanke	Simone	Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Germany
Scholten	Bernice	Netherlands Organisation for Applied Scientific Research (TNO)
Scholz	Stefan	Helmholtz Centre for Environmental Research - UFZ, Germany

Schopfer	Christel	BASF SE
Shindo	Chikako	Ministry of the Environment, Japan
Silva	Maria João	National Institute of Health Doutor Ricardo Jorge (INSA), Portugal
Simon	Eszter	Federal Office for the Environment, Switzerland
Slankster-Schmierer	Eryn	Physicians Committee for Responsible Medicine
Spilioti	Eliana	Benaki Phytopathological Institute (BPI), Greece
Takaki	Koki	Ministry of the Environment, Japan
Teixeira	Joao	Instituto Nacional de Saúde, Portugal
Therkorn	Jennifer	ExxonMobil
Tollefsen	Knut Erik	Norwegian Institute for Water Research (NIVA)
Triebskorn	Rita	Steinbeis Transfer Center, Germany
Tristram	Adrian	ExxonMobil Biomedical Sciences, Inc.
Van Nieuwenhuyse	An	National Health Laboratory, Luxembourg
van Thriel	Christoph	Leibniz Research Centre for Working Environment and Human Factors (IFADO), Germany
Vekic	Ana Maria	Brazilian Ministry of Health
Viegas	Susana	NOVA National School of Public Health, Portugal
Villeneuve	Daniel	U.S. Environmental Protection Agency (EPA)
vom Berg	Colette	EAWAG, Switzerland
Watanabe	Haruna	National Institute for Environmental Studies, Japan
West	William "Jay"	American Chemistry Council, US
Wielsøe	Maria	University of Aarhus, Denmark
Wilks	Martin	University of Basel, Switzerland
Woutersen	Marjolijn	National Institute of Public Health and the Environment (RIVM), Netherlands
Zare-Jeddi	Maryam	Shell

Annex A. Case studies of mixture threshold derivations

A1+A2: Estrogenicity effect biomarker levels for assessing environmental exposures

Estrogenicity has been recognized as a priority endocrine disruption endpoint in chemical, environmental (e.g., water quality assessment) and human safety (e.g., occupational health, cosmetics and food) assessments. Chronic exposure to low concentrations of endocrine disrupting chemicals, chemicals with hormonal properties and their mixtures, can have profound endocrine modulating effects (Escher et al. 2014, Brand et al. 2014 DEMEAU project report). Since the alteration of certain hormonal processes in humans and wildlife was first reported and linked to chemical exposures, estrogenicity is one of the most well-studied endpoints related to endocrine disruption.

The general approach with targeted interpretation levels (e.g., technical effect-based trigger value [T-EBT], reference EBT [R-EBT], provisional EBT [P-EBT], and refined EBT [simply referred to as EBT]; **Table A 1.1**) was also applied to estrogenicity.

In this case study, we present two effect biomarkers of estrogenicity to assess environmental exposure to exogenous estrogenic chemicals: estrogen receptor alpha (ERα) transactivation and vitellogenin induction in male fish (**Table A 1.1**). For the environmental health aspects, experimental data sets were available and targeted interpretation and assessment levels could be derived (see **Table A 1.1**, **Table A 1.2** and **Table A 2.3**). The application of effect-based methods to water quality monitoring and practical considerations in their use are discussed in various studies (e.g., Neale et al. 2023).

Table A 1.1. Generalized approach to deriving proposed effect-based trigger (EBT) values for evaluation of estrogenic effects of chemicals based on either estrogen receptor α (ERα) transactivation or plasma vitellogenin in male fish. Each generalized description is explained in detail below and accompanied to values derived for an exemplary assay (in the case of ERα transactivation) or species (in the case of male plasma vitellogenin).

Environmental Health (ENV)	Level	Meaning if exceeded	Generalized, ERα-transactivation	ERα-CALUX (surface water)	Generalized, male plasma vitellogenin	*Pimephales promelas***
EBT	Refined	Environmental health risk indicated	Threshold with high confidence level accounting for test method specific attributes and differences. (Brion et al. 2018)	0.28 ng E2-EQ/L	Within a factor of 10 of those observed in reproductively mature females.	≈1200 µg/mL
P-EBT	Provisional (P)	Environmental health risk may be indicated, but some uncertainty remains**	Indicative risk screening level that assigns in vivo relevance. (Kase et al. 2018, Könemann et al. 2018)	0.4 ng E2-EQ/L	10-fold greater than the Reference-EBT	≈120 µg/mL
R-EBT	Reference (R)	Exposure above normal surface water effect concentration (e.g., > 95%)	Detectable as significantly different from reference or baseline conditions (i.e., associated with control or reference site(s)).	Surface water dependent	Mean of baseline + two standard deviations (reflecting variability among individuals in the population)	12 µg/mL
T-EBT	Technical (T)	Exposure indicated with an acceptable level of accuracy and precision (e.g., >LOD and >LOQ)	Detection (LOD) and quantification (LOQ) limits that are technically achievable with the bioassay applied. (Könemann et al. 2018, Simon et al. 2022)	LOD: <0.01 ng E2-EQ/L* LOQ: 0.01 - 0.1 ng E2-EQ/L*	Detection limit of the vitellogenin measurement method employed.	2.0 µg/mL****

* 1 L samples were analyzed similarly: first up-concentrated (e.g., often 1 L of water concentrated to 1 mL of extract = 1000x up-concentration) and then diluted in the ERα-CALUX bioassay in studies with 87 European surface water samples (16 samples in Könemann et al. 2018 and 71 samples in Simon et al. 2022).

**Uncertainty arises from the fact that this protection level is based on estimated population-relevant effects. The final level (Refined EBT) will be further refined based on the specificity and sensitivity characteristic of the type of transactivation assay used and may be lower or higher than the proposed P-EBT.

***Based on basal vitellogenin concentrations observed in laboratory reared fathead minnows (*Pimephales promelas*) as documented by Watanabe et al. 2007.

****Ankley et al. 2002, see A2.

A1: *In vitro* Estrogen Receptor alpha (ERα) transactivation and protection levels

Estrogen receptor alpha (ERα) is among the family of nuclear receptors that serve as ligand activated transcription factors. In the case of the ERα, binding of endogenous ligands such as 17β-estradiol (E2), lead to translocation of the ligand-receptor complex to the nucleus where it binds to specific promoter regions within the DNA (estrogen response elements), recruits additional co-activators and/or co-repressors, and subsequently activates or suppresses expression of estrogen-responsive genes. The ability of various xenobiotics, and/or their mixtures, to bind and activate ERα can thus be evaluated using transcriptional activation assays.

In general, ERα transactivation assays detect a highly specific cellular response, such as reporter gene expression, triggered by the receptor binding of (xeno)estrogens. For the example of ERα-CALUX (Estrogen Receptor- **C**hemically **A**ctivated **LU**ciferase gene e**X**pression) method, the activation of the reporter gene leads to the production of an enzyme (luciferase) that yields a quantifiable light signal when a substrate (luciferin) is introduced. This light signal is proportional to the concentration and potency of biologically active (estrogen mimicking) chemicals present in the sample.

ERα activation as a molecular initiating event has been causally linked with adverse effects including reduced fertility and fecundity, altered reproduction (embryo production) with consequent population effects (e.g., Kidd et al. 2007). Not all links between ERα activation and an adverse outcome are fully understood, but there is a compelling weight of evidence linking ERα activation to adverse reproductive and developmental effects in a wide range of vertebrates (e.g. Tang et al. 2020, Gonsioroski et al. 2020, Lucaccioni et al. 2020, Amir et al. 2021). Consequently, ERα transactivation is a useful indicator for identifying chemical exposures that pose hazard as estrogenic endocrine disrupting chemicals.

ERα transactivation can be assessed using well-established *in vitro* test systems. The ISO and OECD validated ERα-CALUX is widely used with the genetically modified human osteosarcoma (U2OS) cell line (ISO 19040-3:2018; OECD TG455 and DB-ALM n°197 JRC-ECVAM automated method protocol). 17-beta-estradiol (E2) is the commonly selected, validated and recommended reference compound (positive control) for ERα transactivation assays. Results are expressed as biological equivalence concentrations (BEQ) of this reference compound (E2-EQ) “that cause the same effect as the effect measured in a sample, a sample dilution or a solution containing one or more test chemicals” (ISO 23196:2022).

We present here the proposed levels of estrogenic effect biomarkers (for ERα-CALUX) for the assessment and evaluation of environmental exposure based on monitoring studies conducted in European surface waters (**Table A1.1**, Könemann et al. 2018, Kase et al. 2018, Brion et al. 2019 and Simon et al. 2022). Effect-based trigger values for a complement of other ERα transactivation assays, derived using the same approach applied to the ERα-CALUX are also provided (see **Table A 1.2**).

(Technical) T-EBT reflects the lowest accurately achievable limit of detection (LOD) or quantification (LOQ) in the selected *in vitro* bioassay. Here we indicate both, the LOD and LOQ ranges reported in the surface water monitoring studies. The range largely depends on the LOD/LOQ determination method (e.g., sample-specific or negative control-based) and whether it was defined for native samples or extracts of environmental (water) samples. These aspects, their pros and cons have been extensively discussed in various studies (e.g., Escher et al. 2018, Könemann et al. 2018, Simon et al. 2022, Neale et al. 2023).

Extraction and preconcentration improve quantification and are proved necessary when the estrogenic potentials of the water samples are (expected to be) low (<LOD or LOQ in the bioassay). A typical pre-concentration (i.e., concentration factor) can be around 1000 (1 L water sample concentrated up into 1 mL extract), but can be much higher with high volume sampling (Schulze et al. 2017).

The demonstrated studies (**Table A 1.2**) defined a sample-specific LOQ determination method. It was defined based on the selected 10% effect level induced by the sample, then extrapolated onto the

reference curve to derive an E2-EQ concentration, which is adjusted to account for any pre-concentration of the sample. LOQ was as low as 0.01-0.1 ng E2-EQ/L in the concentrated extracts. The 10% effect level set for the evaluation of the data was based on previous studies and the ISO 23196 protocol (Calculation of biological equivalence (BEQ) concentrations).

The watch list mechanism of the European Water Framework Directive (WFD) (monitoring of selected emerging contaminants such as estrogenic chemicals) suggests a maximum acceptable method detection limit (MDL or LOD) and corresponds to the substance-specific predicted no-effect level (PNEC). In case of the reference estrogen, E2 it is 0.4 ng/L (EU, Commission Implementing Decision 2018/840; Loos et al. 2018). It also accepts bioassays as detection methods if they meet certain performance criteria and show predictive capacity (EC Document 52022PC0540, 2022). As a rule of thumb for methods the LOQ should be 1/3 of the Environmental Quality Standard (EQS*) or the MDL/LOD as low as PNEC (EU, Commission Implementing Decision 2018/840; Loos et al. 2018). These criteria are met.

Looking at other regulations in use, the European Commission Food Safety Regulations (2017/771) accept the use of bioanalytical methods (e.g*., in vitro* receptor transactivation assays) in addition to chemical analysis to detect the combined effect of dioxins and dioxin-like chemicals in food and feed, expressed as TEQs (toxic equivalents). According to this regulation, "the establishment of the limit of quantification (LOQ) is not an indispensable requirement but the method shall prove that it can differentiate between the blank and the cut-off value." Proving this, the defined bioanalytical reporting level should differ from the procedure blank and be at least three times above the level of the procedure blank. It is also suggested defining this reporting level based on known samples with different matrix compositions and containing typical target chemicals, congener patterns rather than based on sample blank responses or a signal-to-noise ratio of the chemical analytical method.

(Reference) R-EBT: This protection level builds on background effect concentrations that are not linked to adverse outcomes in the environment. These background concentrations are defined at reference sites representing mostly pristine environments with no or low estrogenic loads. Existing monitoring data at various sampling sites, i.e., measured and reported estrogenic levels in environmental waters, where validated sample preparation, quantification and test methods were used. R-EBT can largely differ in waste and surface waters due to numerous factors such as, dilution in the environment, spatial and temporal aspects of data gathering, etc. The studies used to exemplify this protection level were conducted in Europe.

(Provisional) P-EBT: At this level, a provisional, *in vitro*-relevant, indicative threshold value is defined, which makes it possible to discriminate between clean and polluted sites i.e., identify hotspots and to predict possible adverse health effects in aquatic organisms. This value is based on chronic fish effect data with population relevance of 11 fish species and was proposed by EU commission as EQS for E2 after a review process (Loos et al. 2018, SCHER Opinion Paper 2011). This P-EBT could be considered as a generic, highly protective threshold for population-relevant effects in 95% fish species investigated (including reproduction toxicity). The use of environmental quality standards to derive this threshold ensures *in vivo* relevance but is not as specific as the next level of protection (refined EBT for specificity and sensitivity). Therefore, in this case, the P-EBT is higher than the EBT, but it can also change in the opposite direction, depending on the characteristics of the assay.

Refined EBT. A certain degree of variability in the predictive capacity of *in vitro* assays is an inherent characteristic (Wagner et al. 2013). The operating protocols, the cell type used and engineered as well as the relative potencies of the known effect-driving chemicals compared to the reference compound, E2 (in this case) may be different for the available/developed ER transactivation assays. These can influence the effect quantification and result in a slightly different estrogenicity of the very same sample. Furthermore, the direct translation of cellular receptor transactivation into estrogenic effects observed *in vivo*, for example in fish, is not straightforward either. Interspecies differences for the estrogen receptor (ER), toxicokinetics such as adsorption, distribution, metabolism and excretion (ADME) of the chemical

influenced by the actual fitness or life stage of individual animals may lead to weaker, stronger or even non-estrogenic response. This is also the reason why we have made test-specific adjustments to the refined EBT (see **Table A 1.2**).

Therefore, a refined, highly specific threshold (**EBT**) for ERα-CALUX is suggested to be derived, while also accounting for the above-mentioned influencing factors and demonstrating a reliable *in vitro/in vivo* correlation. A smaller set of water samples (n=33) was analyzed using the OECD validated *in vivo* EASZY zebrafish embryo assay to demonstrate the ecotoxicological relevance of *in vitro*-based estrogenicity assessment (Brion et al. 2019). The EASZY assay is a small-scale whole-organism assay that uses transgenic zebrafish cyp19a1b-GFP (Green Fluorescent Protein) embryos to detect chemicals acting through estrogen receptors. The relation between in *vitro* E2-equivalent concentrations (from the ERα-CALUX) and the activity of the samples after exposure of transgenic embryos (in the in vivo EASZY assay) was investigated by linear and logistic regressions with two possible outcomes, activity or inactivity. The regression analyses allowed the determination of an optimal cut-off for the ERα-CALUX, above which *in vivo* response was also observed. With other words, above this refined EBT the risk associated with substances acting through the same mode of action in water samples can be predicted *in vivo* (in zebrafish) by identifying true active (sensitivity) and true inactive samples (specificity) (both close to 100% certainty) (Brion et al. 2019). Samples with high sensitivity (true positive rate in %) indicate an *in vitro* estrogenic activity above the EBT with an *in vivo* estrogenic response. Samples with high specificity (true negative rate in %) indicate an *in vitro* estrogenic activity below the EBT with no *in vivo* estrogenic response.

The same regression exercise was conducted between other currently available *in vitro* ERα transactivation assays and the *in vivo* EASZY assay and summarized in **Table A 1.2**. Important to note that genetically modified cells are used in these assays. For the human estrogen receptor, the environment in the yeast cell is somewhat more artificial than in the mammalian (human) cells. There are also differences in membrane structure, types of enzymes present, cellular uptake and metabolism that can result in different sensitivity to estrogenic compounds (Kunz et al. 2017)

Table A 1.2. In vitro ERα transactivation-specific effect-based trigger values (EBTs) to evaluate estrogenic effects of chemicals or samples (according to Brion et al., 2019).

EBT Type	MELN [ng E2-EQ/L]	ERα-GeneBLAzer [ng E2-EQ/L]	HeLa-9903 [ng E2-EQ/L]	p-YES [ng E2-EQ/L]
Remarks	Mammalian cells	Mammalian cells	Mammalian cells	Yeast cells-based test method, where direct combination with analytic (chromatography and simplified chemical identification) possible
EBT	0.56	0.24	0.18	0.50
Provisional-EBT	0.4	0.4	0.4	0.4
Reference-EBT	Can differ largely depending on surface water situation.			
Technical-EBT	LOD: <0.01 LOQ: 0.07-0.03	LOD: <0.03 LOQ: 0.02-0.11	LOD: <0.04 LOQ: 0.01-0.13	LOD: <0.07 LOQ: 0.01-0.2

** The present values are experimentally defined and reported by Könemann et al. 2018 and Brion et al. 2019. They are not based on a systematic review of all available data in the peer-reviewed literature.*

Closing remarks

This case study, due to the completeness and availability of data at all levels, focused on environmental health. We have seen that an EBTs for the estrogenic endpoint can be slightly variable depending on the assay chosen and the target of protection (e.g., species of interest). Nevertheless, the various EBTs available to date for estrogenic activity in the environment and derived for various assays and in different studies (e.g., Kunz et al. 2015, Könemann et al. 2018, Kase et al. 2018, Brion et al. 2019, Simon et al. 2022) fall within a narrow range of 0.18 and 0.56 ng EEQ/L. This indicates that the range is realistic and should be considered highly acceptable.

A2: Vitellogenin in Male Fish as an effect-biomarker

Vitellogenin is a glycophospholipoprotein that serves as a precursor for yolk production in oviparous vertebrates as well as egg-laying invertebrates (Arukwe and Goksøyr 2003, Tyler and Sumpter 1996; Valle 1993). In oviparous vertebrate animals, vitellogenin synthesis is regulated by estrogen signaling (Wallace, 1985, Romano et al. 2004). During normal reproductive cycles, increases in endocrine or paracrine estrogen signaling lead to increased vitellogenin production in females which is used to supply oocytes with yolk reserves needed to support development (Tyler and Sumpter 1996). However, basal vitellogenin synthesis and circulating abundance in males is very low, even though the vitellogenin gene and subsequent translated protein remain highly inducible in males (Tyler et al. 1996). As a result, vitellogenin induction in male fish has proved to be both a sensitive and highly responsive (in terms of fold-change in abundance) biomarker of exposure to exogenous estrogenic substances (Brown et al. 2023; Dang et al. 2016, Hiramatsu et al. 2006, Solé et al. 2001). These attributes have made plasma vitellogenin concentrations and/or vitellogenin transcript abundance in male fish one of the most widely applied biomarkers in ecotoxicology.

Although vitellogenin induction in male fish is widely accepted as a useful biomarker of exposure to exogenous estrogenic chemicals (Brown et al. 2023, Jones et al. 2000, Sumpter and Jobling 1995), some authors have suggested that vitellogenin induction in male fish is not a strong predictor of adverse effects (Jones et al. 2000). This conclusion was based, in part on the observation that other reproductive endpoints can be more sensitive than vitellogenin induction (Jones et al. 2000). It also reflects a recognition that vitellogenin induction in males is not mechanistically related to adverse effect(s), apart from possible kidney pathology under extreme exposure conditions likely to occur only in laboratory settings (Folmar et al. 2001, Zaroogian et al. 2001). Consequently, while vitellogenin induction after exposure to sublethal concentrations in males is considered a reliable marker of exposure to estrogenic substances, it is not a key event in adverse outcome pathways (AOPs) linking estrogen receptor (ER) agonism to adverse reproductive or developmental effects, as it is not part of the causal series of events. Nonetheless, given the strong link between ER agonism and adverse effects (e.g. Parrott and Blunt 2005, Caldwell et al. 2008, Schubert et al. 2008, Kidd et al. 2014, Matthiessen et al. 2018), and the reliability of vitellogenin induction as a confirmatory measurement of activation of the molecular initiating event (estrogen receptor activation), others have argued that vitellogenin induction in males can be reasonably employed as an effect biomarker, even though it is not part of the causal sequence leading to the adverse effects observed (i.e., excess vitellogenin is not causally linked to decreased fertility in males, altered male reproductive behavior, feminized secondary sex characteristics, etc. which represent effects of estrogens on male fish (Wang et al. 2018, Lahnsteiner et al. 2006, Colman et al. 2009, Parrott and Blunt 2005, Länge et al. 2001). Indeed, Cheek et al. 2001 concluded that vitellogenin induction may be interpreted as an indicator of reproductive hazard, but absence of vitellogenin production should not be interpreted as lack of hazard. Vitellogenin induction may not always be the most sensitive response to estrogenic endocrine disruption (Scholz and Mayer 2008), but the occurrence of significant vitellogenin induction can signal a high probability of adverse estrogenic hazard. Additionally, vitellogenin induction in males typically occurs at concentrations less than or equal to those at which vitellogenin is significantly impacted in females (Brown et al. 2023).

Here we propose a provisional set of guidance for the use of vitellogenin induction in male fish as an effect biomarker. Because the vitellogenin induction itself is not causally linked to adverse outcomes associated with (xeno)estrogen exposure, the effect-based trigger (EBT) values are based on magnitudes of response commonly observed at concentrations of estradiol equivalents (E2-EQ) that have been reported to produce adverse effects. It is noted that baseline vitellogenin concentrations in male fish vary by species (Hiramatsu et al. 2006). Consequently, reference EBT levels intended to represent concentrations that would generally be regarded as significantly different from baseline should be defined for each species of interest (**Table A 1.1**, **Table A 2.3**). Likewise, “Refined-EBT” levels that are intended to be indicative of a strong probability for adverse effects may also be species dependent. In general, concentrations of E2-equivalents that

induce vitellogenin concentrations in males that are similar (within a factor of 10) of those that are typically observed in reproductively mature females are associated with a high probability of eliciting adverse reproductive and/or developmental effects (**Table A 1.1**, **Table A 2.3**). Lesser magnitudes of vitellogenin induction may be associated with adverse effects in some circumstances, but their reliability as a biomarker of effect at lower levels of induction in males may be less clear. Thus "provisional-EBTs" are defined as levels of induction that are at least 10-folder greater than the "reference-EBT", but still one or more orders of magnitude lower than those generally observed in reproductively mature females (**Table A 1.1**, **Table A 2.3**). The use of vitellogenin concentrations in males, as related to concentrations observed in sexually mature females, as the basis for setting the EBTs is empirically-based, considering studies from a range of different species (**Table A 2.3**). However, we note that a comprehensive systematic review was not performed. Once baseline levels of vitellogenin, along with the variability around that baseline (reflecting interindividual variability of the sampled population) are determined (Burden et al. 2023), the general criteria provided (**Table A 1.1**) can be used to define preliminary EBTs.

Even after species-specific EBTs have been defined, caution should be applied in interpreting vitellogenin induction in male fish as an estrogenic effect biomarker in field settings. Baseline vitellogenin concentrations defined for male fish reared in laboratory settings cannot be assumed to be representative of baseline under field conditions (Hara et al. 2016). Likewise, baseline concentrations may vary from one field location to another due to both environmental differences (e.g., temperature, water quality, food quality and availability, etc.) as well as population-specific differences (e.g., genetic or epigenetic variability). Consequently, vitellogenin induction in male fish is best used as an effect biomarker when the reference population used to determine baseline concentrations is closely matched to the experimental population of interest. As such, laboratory exposures or caged fish exposures employing laboratory reared organisms tend to yield less ambiguous results than those obtained with field collected fish of unknown provenance (Hara et al. 2016). However, with these considerations in mind, use of vitellogenin induction in male fish as an effect biomarker should be achievable for a range of fish species (**Table A 2.3**) exposed to variety of estrogenic mixtures, including those that may contain anti-estrogens as well.

Table A 2.3. Preliminary* species-specific effects-based trigger values (EBTs) for the use of plasma vitellogenin concentrations in lab-reared male fish as an indirect effect biomarker for indicating potential for adverse estrogenic effects in fish.

EBT Type	*Pimephales promelas*	*Danio rerio* (protein normalized)	*Danio rerio*	*Oryzias latipes*	*Oncorhynchus mykiss*	*Cyprinus carpio*	*Micropterus salmoides*
Effect-EBT	≈1200 µg/mL	≈2000 µg/mg	60 µg/mL	VTG1 - 13000 µg/mL; VTG2- 1600 µg/mL	1000 µg/mL	≈3000 µg/mL	≈1000 µg/mL
Provisional-EBT	≈120 µg/mL	≈20 µg/mg	2.0 µg/mL	100 µg/mL	50 µg/mL	≈100 µg/mL	≈100 µg/mL
Reference-EBT	12 µg/mL	2.3 µg/mg	0.16 µg/mL	10 µg/mL	5 µg/mL	10 µg/mL	10 µg/mL
Technical-EBT	2.0 µg/mL	0.84±0.73 µg/mg**	0.04 µg/mL	0.02 µg/mL	0.01 µg/mL	1 µg/mL	1 µg/mL
Reference(s)	Watanabe et al. 2007; Ankley et al. 2002	Fenske et al. 2001	Brion et al. 2009	Fujiwara et al. 2005; Scholz et al. 2004; Kordes et al. 2002	Bon et al. 1997	Hiramatsu et al. 2006 ; Lee et al. 2001; Goodbred et al. 2006	Goodbred et al. 2006

* The present values are not based on a systematic review of all available data in the peer-reviewed literature. It is likely that these values could be improved and revised through systematic meta-analysis of hundreds of published studies reporting vitellogenin concentrations in a wide range of fish species.

** Limit of detection (LOD) as reported by Fenske et al. 2001, reflecting some variability in the LOD from assay to assay.

A3: Genotoxicity mixture threshold assessment in Occupational Biomonitoring; the lymphocyte Cytokinesis-block micronucleus (L-CBMN) assay

Summary abstract

There are several assays available to detect genotoxicity in human studies. Out of all options, we selected the micronucleus (MN) assay because of the well-described association to an increased risk of cancer, both via Adverse Outcome Pathways as well as in prospective cohort studies that found significant associations between overall cancer incidence and an increased MN frequency in healthy subjects. This approach can be used to assess genotoxic mixtures. We derived human health risk levels for MN based on the L-CBMN assay. More specifically, for the ROBEL evaluation we propose to establish the baseline MN frequency specific for each laboratory, taking into account major confounders among non-exposed subjects, i.e., age, sex and interlaboratory variability. For the sake of feasibility (quite often the database of one laboratory is small), large age-classes can be used, e.g., ≤50 years; > 50 years. For the OBEL - to identify those subjects for which a health risk is indicated – we propose to assess the micronuclei mean ratio (MN-MR) comparing MN mean frequency in the exposed group with respect to a control group (or historical baseline) frequency of each laboratory, adjusted by sex and by age-class. According to evidence from various large studies and meta-analyses we advise that whenever there is a statistically significant increase of two folds (MN-MR≥2.0) in exposed workers *vs* matched unexposed controls, control measures need to be undertaken to mitigate the risk. We acknowledge the inherent limitations of the L-CBMN assay, and we see promise for other assays to detect MN frequencies, such as in exfoliated buccal cells. However, until other assays are associated to cancer to the same extent as the L-CBMN assay, these other genotoxicity assays can only be used supportive for the risk assessment approach as outlined in this work.

1. Introduction

A genotoxicant refers to a chemical that induces adverse effects on genetic components via any of a variety of mechanisms, including mutation, but does not necessarily connote the ability to cause heritable changes (OECD 2017). The term mutation refers to permanent changes (including gene mutations, structural and numerical chromosomal alterations) in the structure or amount of the genetic material of an organism that can lead to heritable changes in its function (WHO 2020). Genotoxicity testing also includes test for assessing DNA damage, “which may be reversed by DNA repair processes or other known cellular processes or result in cell death and may not result in permanent alterations in the structure or information content of the surviving cell or its progeny” (OECD 2017). An overview with examples of genotoxicity assays is given in **Table 3.8**.

Within this OECD working group the following genotoxic assays were investigated in more detail: i) Cytokinesis-blocked micronucleus (CBMN) assay in human peripheral blood lymphocytes (PBL) (OECD TG 487), ii) Comet assay with frozen samples (OECD TG 489), and iii) Micronucleus assay in Transferrin-positive Reticulocytes (OECD TG 474). To be able to select the most relevant genotoxicity assay, several considerations were taken into account, including the specificity and sensitivity of the marker, to which extent the marker has been used in occupational studies, and how the marker relates to an adverse effect. The micronucleus test is used to identify substances that cause cytogenetic damage which results in the formation of micronuclei (MN). It is widely used in molecular epidemiology and genotoxicity assessment to evaluate the presence and the extent of chromosomal damage in human populations (Bonassi et al. 2007). Scoring MN in binucleated (BN) cells is referred to as the traditional CBMN assay, for which a detailed protocol was published in (Fenech 2007). The *in vitro* micronucleus test used to evaluate chemical genotoxicity is described in OECD TG 487: *In Vitro* Mammalian Cell Micronucleus Test. An important limitation of the assay is the limited sensitivity. In several multi-biomarker biomonitoring studies, CBMN is among the least sensitive assays for the effect of genotoxic exposures (Bonassi et al. 2003). A

chromosomal aberration (CA) is often defined as the appearance of missing, extra, or irregular portions of chromosomal DNA (NHGRI 2016). Although chromosomal aberration assays have been used extensively in the past, these assays are gradually replaced by the *in vitro* MN assay (OECD TG 487). Amongst others because the *in vitro* MN assay is able to detect both clastogenic and aneugenic events (WHO 2020). In addition, the *in vitro* MN assay is faster, less subjective and more advanced automated scoring methods are now available based on image flow cytometry and artificial intelligence to score the micronuclei in BN cells (Wills et al. 2021).

Another popular assay in occupational biomonitoring studies is the Comet assay. This test is able to identify the presence of DNA damage (strand breaks and other lesions that are converted into strand breaks under alkaline conditions) and DNA repair activity. In contrast to the MN assay, this marker is not measuring chromosomal damage but instead it focuses on earlier events for genotoxicity.

Furthermore, contrary to the CBMN assay, the Comet assay does not measure whole chromosome mal-segregation events that lead to aneuploidy (Fenech et al. 2020).

Recognized advantage of this biomarker is the sensitivity to genotoxic exposure, while its weaknesses include technical variability and limited association with health outcomes, as recently addressed by international collaborative efforts, i.e., hCOMET (Møller et al. 2020, Milić et al. 2021, Bonassi et al. 2021).

Here we aim to derive **human health related risk levels** for a robust assay for genetic damage which is well known for assessing genotoxicity in human studies. Therefore, the MN assay is selected, because of the well-described association to an increased risk of cancer, both via Adverse Outcome Pathways (AOPs) (described in more detail in **Chapter 2.1**) as well as published prospective cohort studies that may associate micronuclei frequencies and an increased risk of developing cancer (**Chapter 2.2**). **Chapter 3** provides the derived human health levels, including rationale, for micronuclei frequencies. This report ends with a discussion (**Chapter 4**).

2. Micronuclei in relation to cancer

2.1. Adverse Outcome Pathways (AOPs)

Several AOPs link micronuclei to cancer **(Table A 3.4 and A 3.8**). AOP 293 and 294 both describe how two molecular initiating events (MIEs): DNA damage, and reactive oxygen and nitrogen species (RONS), lead to breast cancer. The authors clarify that increases in DNA damage includes DNA damage in the form of single and double strand breaks, chromosomal damage and micronuclei, as well as some forms of nucleotide damage. AOP296 provides evidence that oxidative DNA damage is leading to chromosomal aberrations and mutations. AOP 272 describes how inadequate DNA repair (as key event (KE2) can lead to an increase in chromosomal aberrations (KE3), and subsequently the development of lung cancer as adverse outcome. Examples of chromosomal aberrations include micronuclei (other chromosomal aberrations listed were abnormal chromosome number (aneuploidy), deletions, translocations, inversions, dicentric chromosomes, nucleoplasmic bridges, nuclear buds, centric rings, and acentric fragments).

In summary, multiple AOPs identify structural and numerical chromosomal aberrations, and micronuclei formation, as KEs towards cancer development. To be able to estimate the extent to which these aberrations can be associated with adverse health effects, there is a need for studies that relate the marker in human tissue and associated cells (e.g. lymphocytes) to the adverse outcome (e.g. cancer). That is the topic of the following subchapter.

Table A 3.4. Overview of Standard Operating Procedures (SOPs) related to DNA damage, chromosomal aberrations and cancer.

AOP#	Name SOP	MIE	KE's	AO	Status
443	DNA damage and mutations leading to Metastatic Breast Cancer	-Increased, DNA damage and mutation	- increased chromosomal aberrations	-metastatic breast cancer	Under development
293	Increased DNA damage leading to increased risk of breast cancer	-Increase, DNA damage	-Increase, Mutations -Increased, Ductal Hyperplasia	-Breast cancer	Under development
294	Increased reactive oxygen and nitrogen species (RONS) leading to increased risk of breast cancer	-Increase in reactive oxygen and nitrogen species (RONS)	-Increase, DNA damage -Increase, Mutations -Increased, Ductal Hyperplasia	-Breast Cancer	Under development
272	Deposition of energy leading to lung cancer	-Deposition of Energy	KE1: DNA Double-Strand Breaks, Increase KE2: Inadequate DNA repair, increase KE3: Mutations, increase KE4: Chromosomal aberrations, increase	-Increase, lung cancer	Endorsed by OECD Working Parties (WPHA/WNT)
296	Oxidative DNA damage leading to chromosomal aberrations and mutations	-Increase, Oxidative damage to DNA		-Increase, Mutations -Increase, Chromosomal aberrations	Endorsed by WPHA/WNT

2.2 Cohort data on micronuclei and cancer

While the evidence generated by mechanistic and basic research is strong and confirms the role of micronucleation in the pathway leading to the diagnosis of diseases, epidemiological data are limited. Most convincing data comes from a historical cohort study within the HUmanMicroNucleus (HUMN) international collaborative project (Bonassi et al. 2007) with a database of 6718 healthy subjects. The subjects came from 10 different countries and were screened in 20 laboratories for MN frequency between 1980 and 2002 in ad hoc studies or routine cytogenetic surveillance, and were subsequently followed up for cancer incidence or mortality. The subjects were chosen because of their exposures to mutagens or carcinogens, or as unexposed referents. The follow-up period began with the date of MN testing and ended with death, cancer diagnosis, emigration, 85th birthday or end of follow-up (1999–2004, depending on the country), whichever occurred first.

Subjects were classified according to the percentiles of MN frequency distribution within each laboratory as low <33%), medium (34 – 66%) or high frequency (>66%), to standardize for the inter-laboratory variability. The mean MN frequencies per 1000 BN cells were provided per country and ranged from 3.9 in Belgium (all males, half of the subjects exposed to potential carcinogens at the time of blood sampling, and half of them were smokers) to 59 in Croatia (71% males, 67% exposed and about half smokers). Significant associations between overall cancer incidence and MN frequencies were found for subjects in the medium (Relative Risk (RR) = 1.84; 95% confidence interval (CI) 1.28–2.66) and high tertiles (RR = 1.53; 95% CI 1.04–2.25), when compared to the low tertiles[1]. With regard to specific causes of death, a

[1] Bonassi and colleagues provided some suitable explanations for the lack of linearity in the relationship between MN frequency and the risk of cancer. The first was based on the assumption that there is a value of MN frequency beyond

statistically significant association was found only in the groups of urogenital and gastro-intestinal cancers. Note that the association between MN frequency and risk of cancer and other diseases is not modified or confounded by factors such as age, gender and occupational exposure to genotoxins and smoking status (Bonassi et al. 2007).

There are also minor nested case-control studies (Murgia et al. 2007, Federici et al. 2008, Murgia et al. 2008), and many case-control studies linking MN frequency to several diseases (summarized in the special issue on 'Micronuclei and Diseases' Mutat Res Rev, 2021).

Results from these studies are consistent and robust, although the role of reverse causality (MN formation due to the disease and not *vice versa*) cannot be ruled out in case-control studies. In addition to these studies, indirect evidence comes from the observation that most carcinogens are genotoxic and induce an increase of MN (Fenech et al. 2016a, Nerseyan et al. 2016), and from validation studies of other DNA damage biomarkers such as chromosomal aberrations.

3. Risk Classes

This chapter provides an overview of the proposed derived human health levels for MN based on the L-CBMN assay.

3.1 OECD project on effect biomarker levels

In 2022, the OECD Occupational Biomonitoring Guidance Document was published as a joint activity of the Working Parties on Exposure and Hazard Assessment involving more than 40 institutes/organisations. The Guidance Document presents current regulatory and scientific approaches to derive occupational biomonitoring values and provides practical guidance on how to use them for risk assessment. In the OECD Guidance Document, the derived health-based human exposure biomarker assessment values are referred to as Occupational Biomonitoring Levels (OBLs) which are suitable for use in risk assessment and risk management. Following this Guidance document, a new OECD effect-biomonitoring project using Adverse Outcome Pathways (AOPs) for mixture assessment was initiated. The aim of this project is amongst others the derivation of practice-oriented guiding principles, including mixture threshold derivation concepts for relevant effect biomarkers.

A summary of the assessment levels can be found in the **Table A 3.5 Chapter 3.2** of this document provides details on the Technical Occupational Biomonitoring Effect Level (TOBEL). The Reference Occupational Biomonitoring Effect Level (ROBEL) is explained in more detail in **Chapter 3.3** and the Occupational Biomonitoring Effect Levels in **Chapter 3.4**. Note that a Provisional Occupational Biomonitoring Effect Level (POBEL) is not proposed in this report as a direct health risk is indicated in relation to several cancer and non-cancer diseases.

which no further increase in cancer risk occurs, e.g. cells with excessive genome damage may be eliminated by apoptosis. Another explanation entails the presence of DNA leakage in the cytoplasm due to an excessive level of MN within cells that may trigger the innate immune response, via the cGAS-STING mechanism, and in turn lead to immunological elimination of some of the cells containing MN [Guo et al. 2023, Beernaert and Parkes 2023].

Table A 3.5. Proposed human health effect levels related to genotoxicity based on the micronuclei mean ratio (MN-MR) in the L-CBMN assay

Human Health (HH)	Level	Meaning if exceeded	Proposal
OBEL (occupational biomonitoring effect level)	Refined	Health risk indicated	Equal or greater than 2 MN-MR - increase in comparison to control group
POBEL	Provisional	Health risk maybe indicated	Not provided
ROBEL	Reference	Exposure above or below reference level (e.g. > 95%, mean	Adjust the control group for age and sex (large age-classes can be used, e.g., ≤50 years; > 50 years)
TOBEL	Technical	Exposure indicated (e.g. >LOQ or >LOD)	MN-MR can always be quantified, without limitation, but depends on cell number count

3.2 TOBEL

The TOBEL relates to the technical achievability to measure a certain biomarker. Since there is no limitation, i.e., a LOD, for measured micronuclei, there is no TOBEL proposed for this specific biomarker.

MN frequency can be measured in peripheral blood lymphocytes collected from human blood samples of exposed population collected in lithium heparin tubes and stored between 5°C and 22°C for up to 24h using the protocol published in 2007 in Nature Protocols (Fenech 2007). The slides containing cytokinesis-blocked binucleated cells can be reliably scored visually using the scoring criteria. Automated scoring, including flow cytometry image analysis has also been shown to be feasible (Fenech 2007, Rossernova et al. 2011, Schunck et al. 2023, Rodrigues et al. 2023).

3.3 ROBEL

The ROBEL relates to reference levels for the biomarker of interest. To derive such a value for micronuclei, the study of Bonassi et al. 2001 is especially relevant. In this study the HUMN project pooled data from a large dataset of participating laboratories, and estimated an overall median MN frequency of 6.5 per thousand in unexposed subjects. The same study recommended a range of three to 12 per thousand as anormal MN frequency in unexposed subjects (Bonassi et al. 2001). However, according to (Fenech 2007), the variability may be even greater, ranging from zero to 30 MN per thousand binucleated cells, depending on experimental and tissue conditions. Apart from occupational, environmental or lifestyle DNA damaging agents, confounding factors that must be considered in evaluating the ROBEL are age, and to a lower extent sex (Bonassi et al. 2001, Fenech 2007, Fenech and Bonassi 2011).

Therefore, considering that the set of major confounders in non-exposed subjects includes age, sex and interlaboratory variability, it is proposed to provide a baseline MN frequency specific to each laboratory, taking into account age-class and sex**.** For the sake of feasibility (when the database of one laboratory is small), large age-classes can be used, e.g., ≤50 years; > 50 years.

We recommended adjusting for sex as this is an important proxy for several variables related to genotoxic exposure, life-style, and diet (Fenech and Bonassi 2011). The choice of 50 years as threshold age is a compromise between 1) mechanistic evidence showing a much higher frequency of MN in older age-classes, 2) epidemiological reports showing that most population studies have a mean age around this age, and 3) statistical possibility, since in many case series the age of 50 divides the study population into groups of similar size, maximizing statistical power. Besides age and sex, another confounding factor to consider is the effect of dietary deficiencies or excesses, as in certain circumstances - such as obesity and folate deficiency - their impact on MN frequency may be substantial (Fenech 2010, Scarpato et al. 2011).

Uncertain is the role of smoking, which appears to play a role in increasing MN only in heavy smokers (Bonassi et al. 2003).

3.4 OBEL

Findings from epidemiological cohort studies cannot be used for individual risk assessment, but pertain to the risk of a group, as also emphasized by (Bonassi et al. 2007). The great inter-laboratory heterogeneity in the measurement of MN prevents the use of quantitative thresholds (such as for cholesterol or glycaemia) to indicate a level of non-acceptable risk. Indeed, validation studies for most biomarkers used group classes for risk estimates, in most cases percentiles, such as in the case of MN, where all subjects were divided within each laboratory by tertile, i.e., low, medium, and high frequency. The approach recommended here - to identify those subjects for which a health risk is indicated (OBEL) - is to assess the Micronuclei mean ratio (MN-MR) with respect to a control group (or historical baseline) frequency of each laboratory of unexposed individuals, stratified by sex and by age-class (e.g., ≤50; 50+). Other lifestyle factors (e.g. exercise, alcohol, smoking and recreational drugs) might influence results as well and need be to assessed in a questionnaire.: A standard sample size estimate should be always planned and only if a proper study is not possible, thumb rules should be applied. For regulatory implications, all studies should be classified as ‘adequately powered’ or ‘underpowered’. This is useful to avoid false negative risk assessments. OHP have to declare in advance what type of statistical analysis they will apply and which results will be evaluated. Note, that we refer here to a comparison between exposed and control groups. Although we refrain from interpreting individual results, we recommend special care if outliers are found. Three risk classes can be evaluated to classify subjects/groups as being at i) non-significant, ii) significant and iii) high risk of developing cancer (see **Table A 3.6**).

Table A 3.6. Relative risk classes s for genotoxicity

Risk class	Definition	Indication
Non-significant	Statistically non- significant difference from non-exposed control group	No measurable (collective) genotoxic exposure
Significant	Statistically significant difference from non-exposed control group, but less than 2.0 x control group induction	Genotoxic exposure health risk cannot be excluded
High	Above 2.0 x induction compared to non-exposed control group	Health risk indicated

A high increase over the expected values calls for intervention with different priority and efficacy depending upon the ratio observed. The **2.0 micronuclei mean ratio (MN-MR)** with respect to the baseline value is based, among others, on a published meta-analysis by (Nersesyan et al. 2016), who found an average a 2.5-fold increase in MN frequencies across different exposure groups (the full dataset included 16959 exposed subjects and 12347 controls) (see **Fig. A 3.5**). The exposure groups included well known genotoxic chemicals such as polycyclic aromatic hydrocarbons and benzene.

Further evidence supporting the 2.0 MN-MR increase is provided by a recent study by (Nersesyan et al. 2024, submitted for publication). This review paper investigated 47 studies measuring MN frequencies in human populations exposed to genotoxic agents in either the occupational setting or as results of environmental pollution (out of 81 studies overall). A total of 13 studies reported a MN mean ratio significantly > 2.0 (27.7%). Interestingly almost all of them concerned occupational exposures, seven of them Formaldehyde (in various groups, and one on Formalin), two on antineoplastic drugs, one each on

lead, pesticides, and arsenic. Only one study resulted positive on environmental exposure, i.e., radiofrequency in mobile phone users. These results support the idea that the number of studies revealing exposures (effects) above 2.0 is reasonable (nearly 30%), that the large majority of situations possibly requiring intervention are in the occupational setting, that all the positive results concern known genotoxic/carcinogenic agents.

Furthermore, the observation is also supported by the observation that MN frequencies in people with a wide range of diseases including cancers are approximately on average 2-fold higher or more compared to healthy controls (**Figure A 3.3**).

The identification of groups at risk requires also a careful evaluation of the contribution given by other sources of genotoxic exposure, such as environmental, dietary and life-style factors, before the exposure can be associated with a significant risk of cancer and other non-cancer diseases (NCD), and direct intervention (at the workplace) or counseling (for individuals) can be implemented.

Nevertheless, in the same way as exposed and unexposed populations are compared with regards to environmental, occupational factors, if there is a statistically significant difference between MN frequency of an exposed group of workers and a matched control group, this result provides a clear indication that control measures need to be undertaken to mitigate the risk.

4. Guidelines

In the past decades, various initiatives have been launched to improve the reporting of scientific research.

For the CBMN assay an extensive protocol was developed by (Fenech 2000). The CBMN assay protocol is described in 46 steps, divided into six main steps: i) isolation and counting of leukocytes, ii) culture of lymphocytes, addition of Cyt-B and harvesting of cells, iii) addition of Cyt-B to culture, iv) harvesting of cells using cytocentrifugation, v) drying, fixing and staining of cells and slide preparation, vi) coverslipping and storage. In addition, clear scoring criteria for different types of cells are presented (Fenech 2000).

For molecular epidemiology studies in general, the STrengthening Reporting of Observational studies in Epidemiology (STROBE) statement was published, which aimed to provide guidance on how to report observational research.

Here we present reporting guidelines for the L-CBMN assay, based on STROBE-ME and complemented with knowledge on properly execution of the assay and known confounding factors which should be taken into account (**Table A 3.7**).

Table A 3.7. Minimum reporting guidelines for the L-CBMN assay, based on STROBE-ME.

Item	Item specification	Adapted STROBE(-ME) guidelines for CBMN assay	CBMN-specific
Methods	Study design	Describe the special study designs for molecular epidemiology (in particular, nested case / control and case / cohort) and how they were implemented.	
	Biological sample collection	Report on the setting of the biological sample collection; amount of sample; nature of collecting procedures; participant conditions; time between sample collection and relevant clinical or physiological endpoints.	
	Biological sample storage	Describe sample storage until biomarker analysis (storage, thawing, manipulation, etc.).	Collect fresh blood by venipuncture into vacutainer blood tubes with lithium heparin anticoagulant. Keep blood tubes at room temperature (i.e., 20–22 1C). Count leukocytes. Resuspend the cells at 1 x 10^6 per mL into the culture medium (Fenech 2000).
	Biological sample processing	Describe sample processing (centrifugation, timing, additives, etc.).	Stimulate mitotic division of lymphocytes by adding phytohaemagglutinin (PHA). Add Cyt-B exactly 44 h after PHA stimulation. Harvest cells using cytocentrifugation. Score at least 1,000 BN cells (Fenech 2000).
	Setting	Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up and data collection.	
	Participants	Report any habit, clinical condition, physiological factor or working or living condition that might affect the characteristics or concentrations of the biomarker.	Essential: gender, age. Preferable: smoking status.
	Variables	Clearly define all outcomes, exposures, predictors, potential confounders and effect modifiers. Give diagnostic criteria, if applicable.	Include, if possible, potential exposure to genotoxic exposures (preferably with measurements).

Item	Item specification	Adapted STROBE(-ME) guidelines for CBMN assay	CBMN-specific
	Data source/ measurement	Laboratory methods: report type of assay used, detection limit, quantity of biological sample used, outliers, timing in the assay procedures (when applicable) and calibration procedures or any standard used.	See method (Fenech 2000).
	Bias	Describe any efforts to address potential sources of bias.	Try to compare similar exposed groups (SEGs).
	Study size	Explain how the study size was determined.	Using a classical power analysis .
	Quantitative variables	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why.	
	Statistical methods	Describe how biomarkers were introduced into statistical models.	
Results	Participants	Report the numbers of individuals at each stage of the study. Give reason for loss of biological samples at each stage.	
	Descriptive data	Give characteristics of study participants (e.g. demographic, clinical and social) and information on exposures and potential confounders	
	Distribution of biomarker measurement	Give the distribution of the biomarker measurement (including mean, median, range and variance).	Calculate the difference (fold increase: MN-MR) of the control group versus the exposed group.
Discussion	Limitations	Describe main limitations in laboratory procedures.	In comparison to (Fenech 2000).
	Interpretation	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies and other relevant evidence.	
	Communication with workers and or employer	Not included in STROBE-ME.	Communicate the group level average (SD) of both exposed and control groups with individual workers and/or employer. Where the results of the assessment referred to in Article 3(2) reveal a risk to workers' health or safety, workers' exposure must be prevented (CMRD, 2004).

Item	Item specification	Adapted STROBE(-ME) guidelines for CBMN assay	CBMN-specific
			If an individual outlier is found: report this outlier back to the individual with the notion that additional care is recommended.

Discussion

Limitations of the L-CBMN assay

The aim of this work was to derive human disease risk levels for the L-CBMN assay. This assay was chosen because it is a robust, widely used assay, and MN has been associated with an increased risk of cancer in a large prospective cohort (e.g. Bonassi et al 2001, 2007, 2008). Despite this there are several limitations of the L-CBMN assay, including its limited sensitivity and a certain extent of technical and biological variability that cannot be completely accounted for. In addition, not all genotoxic compounds induce micronuclei. A possibility could be to additionally assess DNA lesions that are not directly detectable by the L-CBMN assay such as DNA adducts, DNA oxidation (e.g., 8-OHdG) and DNA abasic sites which can be measured using mass spectrometry, histochemical immunoassay and the alkaline comet assay respectively.

However, this suggestion is only relevant in terms of exposure assessment, and not for estimating risk levels; until these markers are associated to cancer to the same extent as the MN assay, these other genotoxicity markers cannot be used yet directly for the risk assessment approach as outlined in **Chapter 3.4**. (In addition, the Comet assay may be confounded by positive results caused by apoptosis or necrosis that may occur following exposure to chemical genotoxins. In contrast the L-CBMN assay does not include necrotic or apoptotic cells when determining MN frequency.)

Exfoliated cells

The L-CBMN assay requires blood sampling. Other, less invasive matrixes for the MN assay including exfoliated cells are currently under consideration. (Bloching et al. 2000) predicted the relative risk of cancer in the upper aerodigestive tract, using the MN-assay in buccal mucosa. (Bolognesi et al. 2021) undertook a systematic review and meta-analysis for head and neck cancer, and breast cancer (BC), in relation to MN test results in both peripheral lymphocytes and buccal cells. The authors found among others significant differences between BC patients (n = 183) and controls (n = 165) and MN in buccal cells. In a very recent review, an excellent correlation was found between MN mean ratio estimates in lymphocytes and in buccal exfoliated cells ($r = 0.560$; $p<0.01$) (Nersesyan et al. 2024, submitted for publication).

An extensive overview of the association between MN in exfoliated cells and major diseases is reported in (Fenech et al. 2021). A recent evaluation of published literature shows that most researchers opt for the MN assay in exfoliated buccal cells given the evident advantage of a non-invasive sampling, easier storage and preservation (Fenech et al. 2021; Fenech et al. in preparation).

As also elaborated in **Chapter 1 and 2**, the mechanistic evidence linking the process of micronucleation to the pathogenesis of cancer measured in PBLs is more exhaustive when compared to the exfoliated cell. Nevertheless, despite the different solidity of the mechanistic background, the level of DNA damage, when simultaneously measured in PBL and buccal exfoliated cells, was highly correlated in the two tissues (Ceppi et al. 2010). This evidence paves the way to the use of MN in exfoliated cells as a biomarker of risk, based on the assumption that similar damage corresponds to similar risk of diseases. However, until

more larger and prospective studies become available for the MN frequency in exfoliated buccal cells, the L-CMBN assay remains our first choice.

Individual risk assessment

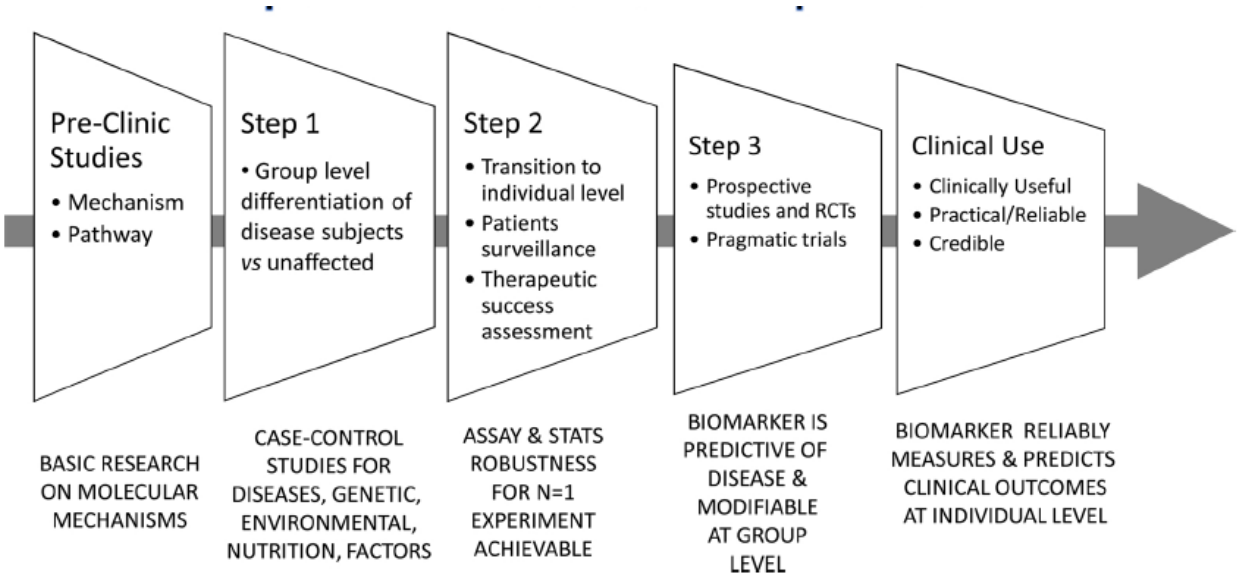
The risk classes are currently based on similar exposure groups (SEGs).

Traditionally, the technical and inter-individual variability of the assays, together with the long duration between early pathogenetic events and the disease, prevented applications to these biomarkers at individual level. These limitations became less stringent with the increased availability of mechanistic and epidemiological data linking MN frequency to the risk of various diseases or clinical altered conditions.

MN provides a level of genetic health of the individuals investigated, and this information, depending upon the strength of the association with disease, may be used for individual risk assessment of for clinical purposes. The history of biomarkers such as cholesterol, which was validated for individual risk assessment thanks to large prospective studies and randomized controlled trials, clearly delineates the roadmap to validate MN beyond group level risk assessment.

The current status of the MN assay in its validation process for individual risk assessment has been extensively discussed by (Bonassi and Fenech 2021) in a paper delineating the roadmap for the validation of this biomarker. **Figure A3.2 and A3.6** illustrate the roadmap proposed to achieve validation and the state of validation of the MN assay regarding its association with some diseases or clinical conditions more frequently investigated.

Figure A 3.2. The validation process of early effect biomarkers to individual risk assessment/use in clinical practice (Bonassi and Fenech 2021).



Other disease endpoints

Extensive evidence exists from cross-sectional studies showing increased MN frequencies in several other diseases, including major chronic diseases. Prospective studies have shown that elevated MN in lymphocytes predicts cardiovascular disease mortality (Murgia et al. 2007, Federici et al. 2008). These trends are also observed in case-control studies over a wide range of cancers and non-cancer diseases. **Figure A 3.3** summarizes the outcomes of meta-analyses published in a special issue (SI) of Mutation

Research on the topic of “Micronuclei and Disease” (Fenech et al. 2021). This meta-analysis supports a broader relevance of MN induction for other health endpoints.

However, the extrapolation of risk thresholds to non-cancer diseases (NCD) is harder to figure out, since reliable prospective data are partially missing. Cross-sectional studies comparing subjects affected by several diseases vs healthy controls found higher MN frequency quite often, and despite the lack of prospective studies, a risk for major chronic diseases cannot be ruled out. The mechanistic association of MN with cardiovascular disease (CVD) and other NCD is not yet known. However, it is plausible because several risk factors of CVD and other NCD such as inflammation, high homocysteine, high advanced glycation end products, methylglyoxal and deficiency in micronutrients such as folate, and vitamin B12 exert genotoxic effects and are associated with increased MN frequency in lymphocytes (Kirsch et al. 2020, Stopper et al. 2004, Donnelan et al. 2020, Fenech M 2012). Furthermore, MN cause inflammation because leakage of DNA from MN into the cytoplasm triggers the pro-inflammatory innate immune response via the cGAS-STING mechanism (Guo et al. 2022, Mackenzie et al. 2017).

Considering methodological weaknesses, and the difficulty to make meaningful recommendations for so many diseases the proposal is that increased MN frequency should raise the awareness in workplaces also for most common chronic diseases and commit employers to control genotoxic exposure and educate workers about the genotoxic risks of poor diet and lifestyle.

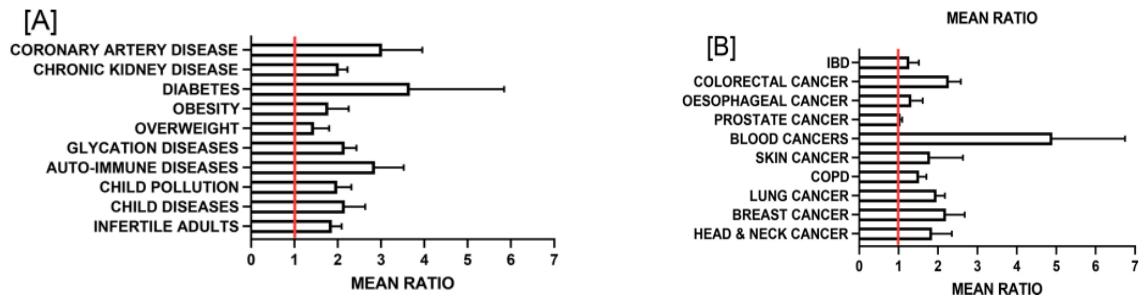


Figure A 3.3. MN ratios associated with non-cancer and cancer diseases (taken from Fenech et al. 2021)

A3 Appendix I

Table A 3.8. Examples of assays for genotoxicity. Adapted from WHO 2020. Red circles refer to the assays which were considered by this OECD working group

Gene mutation	Chromosomal damage	DNA damage/repair
In vitro assays		
Bacterial tests [see section 4.5.2.1] • Reversion to a specific nutrient independence in *Salmonella typhimurium* and *Escherichia coli* (OECD TG 471) *Mammalian tests [see section 4.5.2.2]* • Forward mutation at the *TK/Tk* gene (OECD TG 490) in cell lines such as mouse lymphoma L5178Y and human TK6 • Forward mutation at the *Hprt/HPRT* gene (OECD TG 476) in primary cells or cell lines such as mouse lymphoma (L5178Y), Chinese hamster ovary (CHO), Chinese hamster lung (V79), human TK6 and human lymphocytes	• Sister chromatid exchange (OECD TG 479)[a] • Chromosomal aberrations (OECD TG 473) in CHO, CHL or V79 cell lines and human cells (lymphocytes and TK6) [see section 4.5.2.4(a)] • MN (resulting from clastogenicity and aneuploidy) (OECD TG 487) in CHO, CHL or V79 cell lines and human cells (lymphocytes and TK6) [see section 4.5.2.4(b)] • Chromosomal aberrations (OECD TG 490) in mouse lymphoma L5178Y and human TK6 cells [see section 4.5.2.4(c)]	• UDS in primary cultures (often hepatocytes; OECD TG 482)[a] • DNA strand breakage and alkali-labile sites monitored by single-cell gel electrophoresis (comet assay) or by sucrose gradient, filter elution or alkaline unwinding, in cell cultures [see section 4.5.2.6] • Upregulation or stabilization of DNA damage responses (e.g. p53, ATAD5, pH2AX) • DNA adduct measurement in cell cultures
In vivo assays		
Somatic cell assays [see section 4.5.2.3(a)] • Transgenic rodent assays: *gpt*, Spi$^-$ (*gpt* delta mouse or rat), *lacZ* plasmid, bacteriophage or *cII* (Muta™Mouse) or *lacI* or *cII* (Big Blue® mouse or rat) (OECD TG 488) • *Pig-a* gene mutation assay (mouse, rat, human) *Germ cell assays [see section 4.5.2.3(b)]* • Specific locus test (mouse) • Dominant lethal assay (rodents) (OECD TG 478) • Transgenic rodent assays: *gpt*, Spi$^-$ (*gpt* delta mouse or rat), *lacZ* or *cII* (Muta™Mouse) or *lacI* or *cII* (Big Blue® mouse or rat) (OECD TG 488)	*Somatic cell assays* • Sister chromatid exchange (OECD TG 482)[a] in bone marrow (rodent) • Chromosomal aberrations (OECD TG 475) [see section 4.5.2.5(a)] • MN (resulting from clastogenicity and aneuploidy) (OECD TG 474) in erythrocytes (rodent) [see section 4.5.2.5(b)] *Germ cell assays* • Chromosomal aberrations (OECD TG 483) (rodent) [see section 4.5.2.5(a)] • Dominant lethal mutations (OECD TG 478) (rodent)	• Strand breakage and alkali-labile sites monitored by single-cell gel electrophoresis (comet assay) in nuclear DNA in various tissues (OECD TG 489) [see section 4.5.2.7(a)] • DNA adduct measurement [see section 4.5.2.7(b)] • UDS (liver; OECD TG 486) [see section 4.5.2.7(c)]

CHL: Chinese hamster lung; CHO: Chinese hamster ovary; DNA: deoxyribonucleic acid; *gpt*: glutamic–pyruvic transaminase; *Hprt*: hypoxanthine–guanine phosphoribosyl transferase; MN: micronuclei; OECD: Organisation for Economic Co-operation and Development; TG: Test Guideline; *Tk*: thymidine kinase; UDS: unscheduled DNA synthesis

[a] OECD TGs for these assays were deleted in 2014; legacy data may be used in a comprehensive assessment of genotoxicity, but new tests of this nature should not be conducted.

A3 Appendix II

Human Health (HH)	Environmental Health (ENV)	Tier	Level	Meaning if exceeded
OBEL =Occupational Biomonitoring Effect Level	EBT =Effect–Based Trigger value	3	Refined	Health risk indicated
POBEL	P-EBT	2	Provisional	Health risk maybe indicated
ROBEL	R-EBT	1	Reference	Exposure above or below reference level (e.g. > 95%, <5%)
TOBEL	T-EBT	0	Technical	Exposure indicated (e.g. >LOQ or >LOD)

Figure A 3.4. Proposed concept of Occupational Biomonitoring Effect Level (OBEL), Provisional OBEL, Reference OBEL. Technical OBEL allowing their interpretation. LOD= Limit of Detection, LOQ= Limit of Quantification

A3 Appendix III

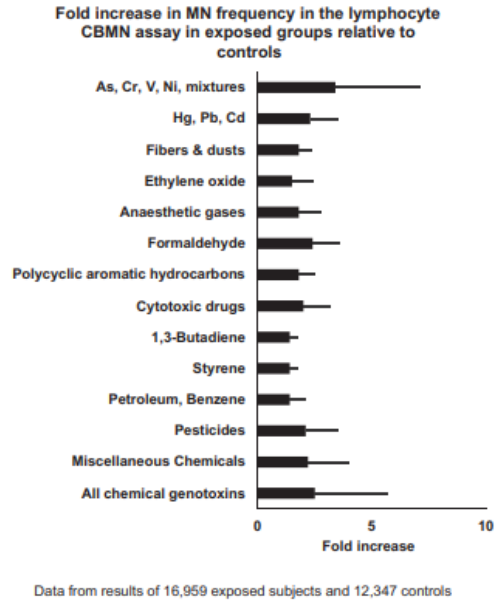


Figure A 3.5. Results obtained in different exposure groups (taken from Nerseyan et al. 2016)

A3 Appendix IV

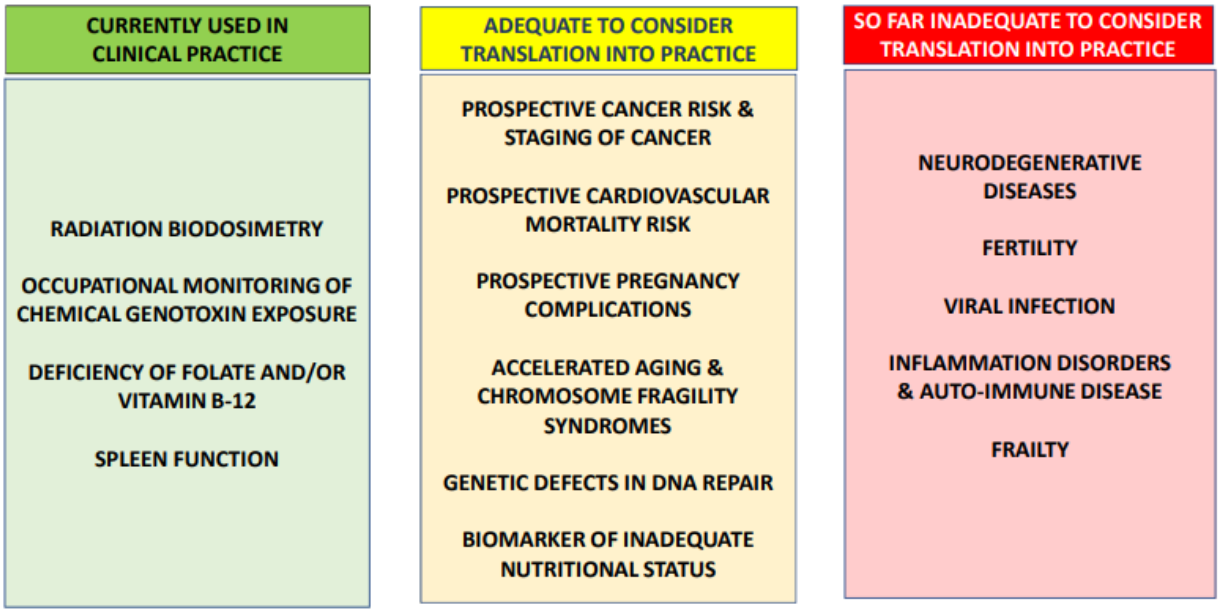


Figure A 3.6. The validation status of MN assay as a clinical biomarker based on recent literature reviews. These reviews provide further information on the validation status of MN in regard to its association with diseases listed in this figure (Bonassi and Fenech 2021).

A4-A6: Neurotoxicity (NT) & Developmental Neurotoxicity (DNT) effect biomarker assessment levels for assessing occupational exposures

The occupational biomonitoring assessment levels are described in the main guiding principles and are applied in the following table for three effect biomarkers:

- Brain Derived Neurotrophic Factor (BDNF) assessment in serum (derivation in A4)
- Neurofilament light chain (NfL) assessment in serum (derivation in A5)
- Neurogranin (NGRN, Ng) assessment in serum (derivation in A6)

Table A 4.9. Occupational assessment levels for (D)NT effect biomarkers

Human Health (HH)	Level	Meaning if changed	A4 BDNF in serum	A5 Neurofilament light chain (NfL)	A6 Neurogranin (NRGN, Ng)
OBEL =Occupational Biomonitoring Effect Level	**Refined**	Health risk indicated	Quantitative relationship of BDNF to adverse effects of higher incidence of the AO in the relevant AOPs is described only in animal or *in vitro* studies. Despite this, several epidemiological studies of lead exposure and cognitive impairment are available (see A4 Appendix Table A 4.11). BDNF levels under occupational exposure are missing.	Several quantitative relationships of NfL concentrations with neurological disease severity (e.g. MS, TBI, anti-cancer drugs) and cognitive performance are available.	Ng not measured in occupational epidemiological studies. Clinical studies reported that Ng in CSF and plasma or serum can be used as biomarker of synaptic dysfunction and cognitive impairment (Liu et al. 2020). Occupational studies with biomonitoring data of neurotoxic chemicals and serum Ng levels are needed to derive OBEL.
POBEL	**Provisional**	Health risk maybe indicated	NT: 150 µg Pb/L equivalents in workers (based on recent 'German BAT value'; OBL). For DNT the levels might be lower. →direct Pb vs BDNF correlation in workers needed. This value is not protective for offspring of female lead-exposed workers at childbearing age.	NT: 150 µg/L Pb equivalents in male workers, correlation analysis ongoing. Based on (Benkert et al. 2022) NfL levels above age specific 95% can lead to reduced	NT: 150 µg/L Pb equivalents in male workers. This value is not protective for offspring of female Pb-exposed workers at childbearing age (link). A correlation between blood Pb and plasma

			As intermediate POBEL a BDNF level of 16.5 µg/L in serum can be used.	structural integrity of the brain: age classes→ NfL (95% percentile [ng/L]) 41-50--> 12-15 51-60--> 15-20 61-70--> 20-26 71-80--> 26-32	Ng in workers is needed.
ROBEL	**Reference**	Exposure not in reference level (e.g. <5%, Mean, >95%)	In adults, mean total BDNF ranges from 32.3 to 40 µg/L. In adolescents it ranges from 26.8 to 31.5 µg/L (see A4 Appendix Table A 4.10); BDNF DNA methylation: no population studies in adults, only in adolescents. For BDNF two human population studies have shown that the mean BDNF serum levels, BDNF= 33 ± 17 µg/L in Latino population in North America (n=349; Fonseca-Partilla et al. 2019) and BDNF= 33 ± 8.5 µg/L in Northern Europe (n=259; Naegelin et al. 2018) are comparable. Therefore, as a ROBEL a mean BDNF level of 33 µg/L in serum can be used.	Due to the extensive clinical use large datasets are already available providing normative data (e.g. z-scores) that could be used as reference level. A fixed cutoff level would be >10 pg/mL measured in serum (Benkert et al. 2022).	In adults (non-diseased controls), serum Ng is usually below 200 µg/l (see A6 Appendix Table A 6.13).
TOBEL	**Technical**	Exposure to neurotoxic substances might be possible (e.g. >LOQ or >LOD)	The LOQ for BDNF is between 7.8 and 62.5 ng/L, with Simoa technology a LOQ of 0.03 ng/L is reached → being detectable in 100% of human samples	LOQ of 0.02 ng/L NfL	LOQ of 0.05 ng/L Ng (ELISA commercial kit)

A4: Brain Derived Neurotrophic Factor (BDNF) assessment in serum

Alteration of Brain Derived Neurotrophic Factor (BDNF) is defined as a Key Event (KE) in a few currently available DNT AOPs since it plays fundamental roles in brain development (neuronal survival, differentiation as well as morphological and functional maturation including brain plasticity). Similar to other neurotrophins, BDNF is first synthesized as a precursor protein (pro-BDNF) which is converted to mature BDNF by extracellular proteases (matrix metalloproteinase-9). Mature BDNF and pro-BDNF elicit opposing effects. Mature BDNF preferentially binds to the tyrosine kinase B (TrkB) receptor and plays an important role through BDNF-TrkB signalling pro-survival pathways. In contrast, elevated pro-BDNF preferentially binds to the p75 neurotrophin receptor (p75NTR) and elicit apoptosis rather than cell survival (Martinowich et al. 2007). An altered balance of the different forms has been linked with cognitive impairment and psychiatric disorders (Carlino et al. 2011, 2013, Garcia et al. 2012).

Both mature and pro-BDNF as well as total BDNF levels can be measured in human blood using commercially available BDNF ELISA kits (Polacchini et al. 2015). It is also measured in neuronal/glial mixed cell cultures at the protein (Pistollato et al. 2020; Egan et al. 2003) and mRNA levels (Zafra et al. 1992, Legutko et al. 2001) or in the cerebral-spinal fluid (CSF) as biomarker of learning and memory impairment in adults (Breno et al. 2014; Yalachkov et al. 2023). The majority of the BDNF transcripts can also be detected in the blood cells, blood serum (Naegelin et al.2018, Polacchini et al. 2015) or plasma (Gejl et al. 2015) and urine (Olivas-Martinez et al. 2023). The peripheral BDNF levels (gene expression and protein) are used as potential biomarkers for neuro-behavioural performance impairment, learning and memory deficit in children and adults, psychiatric disorders such as depression (Polyakova et al. 2015), Parkinson (Rahmani et al. 2019) and Alzheimer's disease (Du et al. 2018, Gao et al. 2022). In addition, based on epidemiological and clinical studies, BDNF DNA methylation is proposed as a highly reliable effect-biomarker since it is more stable over time compared to BDNF gene expression or protein levels (Mustieles et al. 2020 and Mustieles et al. 2022). More information is available in the **Annex B5**.

Justification for TOBEL:

Total BDNF protein levels in human samples

All commercially available tests can detect and quantify serum BDNF protein levels among nearly 100% of the human samples assessed. In general, the analytical LOQ for serum BDNF protein levels is between 7.8 and 62.5 pg/mL depending on the specific assay. Serum BDNF protein levels can also be assessed with digital single molecule array (Simoa technology), showing a LOQ of 0.03 pg/mL or magnetic Luminex Assay Kit for BDNF.

The performance and reliability/ robustness of different commercial ELISA kits to measure serum BDNF in 40 adult serum samples were compared by Polacchini et al. (2015). In this study, ELISA kits from Biosensis for total BDNF and Aviscera-Bioscience for mature BDNF were identified as the most reproducible for measurements of serum BDNF out of six commercially available kits. Biosensis, showed the lowest coefficient of variation even after the third replica (10%) and no significant difference between three independent measurements. Aviscera-Bioscience assays showed acceptable values of 12% CV (Polacchini et al. 2015).

In all cases, the sensitivities of the assays are well below the expected values in humans that showed a mean (SD) of 32.7 ± 8.3 ng/mL in 259 adults and range between 15.8– 79.8 ng/mL (Naegelin et al. 2018) and a median (P25, P75) of 31.5 (25.4, 38.8) and range of 17.2-56 ng/mL among 130 adolescents aged 15-17 (Mustieles et al., 2022).

Based on the available information, there are several reasons indicating that serum should be preferably used to measure BDNF levels compared with plasma or whole blood (Polacchini et al. 2015, Naegelin et al. 2018).

BDNF concentration in serum is about 100-fold higher than plasma levels. Furthermore, BDNF concentration in plasma is affected by handling of the blood sample because of the presence of platelets, which store and can secrete BDNF. Also worth mentioning that the mean levels of BDNF in serum were stable over a period of 12 months as shown by Naegelin et al. 2018.

Therefore, the BDNF quantification in plasma can be sensitive to preparation procedures and is very difficult to be reproducible among different operators. Even when using serum, special attention should be dedicated to the processing of blood, making sure that all samples achieve a complete level of coagulation before serum is collected and analyzed for BDNF (Polacchini et al. 2015).

Current research is focused on: 1) Validating if the mature and pro-BDNF isoforms can be reliably assessed in serum; 2) If measurement of BDNF biomarkers in neural cell-derived exosomes can improve the sensitivity for neurobehavioral effects (Huo et al. 2021).

BDNF DNA methylation levels

Higher BDNF % methylation levels are known to reduce BNDF expression. In humans, the BDNF gene presents a low absolute methylation status. This is the main reason why microarray-based DNA methylome measurements are not a reliable method for its assessment (Forest et al. 2018, Sugden et al. 2020). In this case, bisulfite pyrosequencing would be the gold standard to quantify the percentage of methylation of the CpG islands of the promoter BDNF region of interest (Kundakovic et al. 2014). A LOD or LOQ cannot be established for molecular biology techniques. In general, the percentage of BDNF DNA methylation should be quantified in nearly 100% of the samples (Mustieles et al. 2022).

Justification for ROBEL:

Serum BDNF protein levels

Serum BDNF protein levels in normal population can vary due to both sociodemographic and technical variables. Thus, correction and/or adjustment should be performed considering age, sex, time of blood withdrawal, storage, socio-demographic determinants, and other confounding factors including urbanicity, gender, age, body mass index, smoking status, diet, alcohol intake (Bus et al. 2011) and physical activity (Muñoz Ospina et al. 2024).

In adults, mean total serum BDNF ranges from 32.3 to 40 µg/L and in adolescents, mean total serum BDNF ranges from 26.8 to 31.5 µg/L (see **Table A 4.10**).

Reliable serum BDNF levels, measured in a narrow range in population are available in the epidemiological studies. This information allows a good identification of occupational exposures leading to BDNF serum level alterations.

Two human population studies have shown that the mean BDNF serum levels are comparable: BDNF= 33 ± 17 µg/L in Latino population in North America (n=349; Fonseca-Partilla et al. 2019) and BDNF= 33 ± 8.5 µg/L in Northern Europe (n=259; Naegelin et al. 2018). There were no significant sex differences reported, only a slight age dependency. The studies suggested similar BDNF levels in the population between different continents and in different ethnic groups. However, the differences in standard deviation can be explained by integrating and considering a wide range of age classes (18-85 years with a mean age of 42 years vs. 18-70 years with a mean age of 44 years). Therefore, as a ROBEL, a mean BDNF level of 33 µg/L in serum can be used. Due to the large sample age range in the general populations studied (see **Table A4.9**), the relative standard deviations are expected to be much larger than in control groups for workers, so it would be misleading to make use of unspecific standard deviations. We recommend comparing workers with an unexposed control group of a comparable age class to identify occupational exposures more accurately. One of the further needs is to have age-specific BDNF measurements to identify occupational exposures with higher accuracies.

BDNF DNA methylation

Higher BDNF % methylation levels are known to reduce BNDF expression. In humans, the BDNF gene presents a low absolute methylation status. In adults, there are no general population studies assessing BDNF DNA methylation using pyrosequencing. The range of measurement among 130 general population adolescents aged 15-17 years was 2.7%-5.54%, showing a median (P25, P75) of 3.7% (3.45%, 4.04%) (Mustieles et al. 2022). A new follow-up of these Southern-Spain adolescents is planned for year 2023 when they will be young adults aged around 23 years. For 2024, we expect to have BDNF DNA methylation data for young men adults.

Currently, BDNF methylation has not yet enough datasets in adults to be used in occupationally exposed adults. However, preliminary data suggest that BDNF DNA methylation holds promise as a sensitive effect biomarker and therefore its potential use should be revisited in the future.

Justification for POBEL

Pb induced toxicity is well documented and neurotoxicity (NT) is considered to be the most sensitive endpoint (Greiner et al. 2022). Numerous studies have shown a link between Pb exposure and neurotoxicity in adults, children as well as during pre- and post-natal exposure (developmental neurotoxicity, DNT). Therefore, Pb is used as the reference compound to correlate Pb-induced neurotoxicity with changes in BDNF levels. The mechanistic understanding for BDNF in association with Pb exposure is well understood in animal studies as documented in AOPs 12 and 13 in AOP-Wiki Moreover, epidemiologic studies showing lead as potent prototypical stressor (see **Table A 4.2**) in different life stages of humans in association with adverse neurotoxicity endpoints.

The neurotoxicity of Pb exposure is well documented in children and in the case of prenatal exposure, there is no blood Pb level that can be considered safe. A reference range value was selected based on the 98th percentile of the National Health and Nutrition Examination Survey (NHANES) blood lead distribution in women of childbearing age (Cantor et al. 2019). Most governmental agencies, including the US Environmental Protection Agency and Centres for Disease Control and Prevention (2012), have concluded that based on the epidemiological studies, there is sufficient evidence for adverse health effects in children and adults at blood Pb levels even below 50 µg/L (Chiodo et al. 2007, Bellinger 2008, Jusko et al. 2008). Indeed, there is a reliable number of epidemiological studies in children showing that low-dose Pb levels are associated with low intelligence, anxiety and depression symptoms, among other neurobehavioral impairments (Zhou et al. 2019; Ren et al. 2016; Malavika et al. 2021, Zhang et al. 2017). Additionally, data provided by Malavika et al. 2021 suggests that serum BDNF levels could be responsible for the effect of blood Pb levels on children neurobehavioral alterations.

Based on the recent German BAT derivation including nine neurotoxicity studies, the biological tolerance value (BAT value) for the blood concentration of Pb at the workplace has been defined at 150 µg Pb/L (Greiner et al., 2022) and at this Bioanalytical Equivalent Concentration (BEQ) a NT risk cannot be excluded for male workers. It should be noted that Pb was assigned to Pregnancy Risk Group A (derived by German MAK-BAT Commission), meaning that damage to the embryo or foetus in humans has been unequivocally demonstrated and is to be expected even when the BAT value of 150 µg/L blood is observed. Moreover, in occupationally exposed adults, neurocognitive effects have been reported at blood Pb levels as low as 200–300 µg/L (Mantere et al. 1984, Schwartz et al. 2001) with overt encephalopathy, seizures, and peripheral neuropathy generally occurring at much higher levels (e.g., higher than 1000–2000 µg/L) (Centers for Disease Control and Prevention, 2017). It is well documented in numerous studies that occupational exposure to Pb impacts neurobehavioral performance and causes cognitive impairment (many examples see **Table A 4.11**) including male workers but only Pb levels were measured, not BDNF).

Currently, there are no studies assessing occupational exposure to Pb and measured changes in BDNF levels in workers BDNF can become a valid effect marker for chemically-induced neurotoxicity at the work place, the current assessment knowledge only allows to identify potential health risks.

An occupational exposure study among manganese-exposed smelters assessed plasma BDNF levels and cognitive function (Zou et al. 2014). In this study a dose–response relationship was established between manganese exposure and decreased plasma BDNF levels (control group: 288.7 ± 181.7 pg/mL, low-exposure group: 223.4 ± 178.3 pg/mL (23% mean BDNF reduction), intermediate-exposure group: 178.2 ± 138.1 pg/mL (38% mean BDNF reduction) and high-exposure group: 127.5±99.8 pg/mL (56% mean BDNF reduction). The decreased BDNF levels were determined in 248 selected exposed workers and compared to 100 control subjects, and the linear trend of plasma BDNF levels in different exposure groups was statistically significant ($p<0.01$), after controlling for age and alcohol drinking status. These decreased BDNF levels were also predictive of cognitive impairments evaluated based on the Montreal Cognitive Assessment (MoCA) test ($r=0.278$, $p<0.01$). Compared to the control group that showed a MoCA score of 25.62 ± 0.25 points, the exposure groups showed dose-dependent reductions in cognitive performance: low-exposure group (MoCA = 23.57 ± 0.23); intermediate-exposure group (MoCA = 23.22 ± 0.30); and high-exposure group (MoCA = 21.33 ± 0.32 points). Overall, this important occupational study highlights that: 1) Manganese exposure is dose-dependently associated with a reduction in plasma BDNF levels; and 2) The reduction in BDNF levels correlates with a decrease in cognitive performance as assessed by the MoCA score.

The scientific peer-review literature was searched to identify animal studies linking chemical exposure to Pb, internal concentration in serum and brain BDNF levels, to adverse effects. Three animal studies (Li et al. 2024, Wei et al. 2024 and Wu et al. 2020) were identified where relative BDNF decreases under chemical exposures were observed. Briefly, in the study of Li et al 2024 mice (4–5 weeks of age) were exposed to Pb chloride, resulting in an internal blood Pb effect-Level: 140 µg/L causing a reduction in BDNF mRNA and protein levels in hippocampus of more than 50% (see Fig. 6c in Li et al. 2024). This caused memory deficiency and anxiety-like behaviours in mice.

In the study of Wu et al. 2020 28 postnatal days (PND) offspring mice were exposed to Pb in diet for three months and the level of BDNF in hippocampus was measured at 4-, 13-, 16-months of age, as well as spatial learning and memory function were assessed. The study showed after 4 months a 30% decrease in mature BDNF levels compared to controls and more than 50% decrease after 16 months from birth (see Fig. 5d in Wu et al. 2020). At all exposure intervals consistent deficits of cognition across subsequent age-range were observed (see Fig. 2 and 3 in Wu et al. 2020).

In the study of Wei et al., 2024 mice, aged 8 weeks were exposed to Pb or and Mn in drinking water for 8 weeks resulting in around 30% reduction of BDNF levels and learning and memory impairment. In the case of Pb exposure adverse effects of impaired learning and memory function were observed in mice studies at Pb levels which corresponded to the range of NOAELs for Pb induced neurotoxicity in humans at 180 µg/L. Animals showed after Pb or Mn exposure around 30% reduction of BDNF levels. The combined exposure (to both Pb and Mn) resulted in expected stronger effects, causing around 50% reduction of BDNF levels (see Fig. 4a+b in Wei et al 2024). All three different exposures showed adverse effects however, simultaneous exposure to Mn and Pb resulted in more severe hippocampus-dependent learning and memory impairment. The endpoints in the studies (Li et al. 2024, Wei et al. 2024 and Wu et al. 2020), are comparable in the terms of adversity and in quantitative association with BDNF levels: impairment of learning and memory was observed after a 30-50% reduction in BDNF levels. These results are in line with the occupational worker study of Zou et al. 2024, in which decreased BDNF levels were associated with cognitive impairment and decreased plasma BDNF levels of more than 50% between the group with a highest exposure to Mn and the control group. As mentioned above, the low, intermediate and high exposure occupational groups showed approximated reductions in mean plasma BDNF levels of 23%, 38% and 56%, respectively.

Combining the evidence from the three animal studies and the worker study conducted by Zou et al. (2024), it was observed that a relative reduction of 30-50% BDNF levels was associated with adverse effects such as impaired learning and memory function at internal Pb and Mn levels relevant to workers. Therefore, it is suggested to adopt a conservative approach by considering a 50% decrease in serum of BDNF levels

as an intermediate POBEL compared to a control group to indicate likely adverse effects. This corresponds to approximately 16.5 µg/L BDNF in serum (see Naegelin et al. 2018, Fonseca-Partilla et al. 2019).

The current assessment concept for using effect biomarkers should consider the expected uncertainty and variability of findings. Regarding the uncertainty in the POBEL derivation process, we found severe effects already at reductions of BDNF levels between 20% to 50% in different studies during internal lead exposures comparable to the NOAEL of human studies (≈ 180 µg/L Pb in blood, see Greiner et al. 2022). All key studies were assessed with sufficient reliability and relevance (confidence) and were combined in a weight of evidence approach. Additional certainty is provided by Klein et al. 2011 indicating that blood BDNF concentrations reflect brain-tissue BDNF levels across species. Using a 30% reduction of BDNF as indicator for likely adverse effects would be the most protective. However, considering age-related variabilities of human BDNF levels, a 50% reduction in mean BDNF levels compared to control group would provide additional certainty in identifying strong neurotoxic effects. In order to avoid an overinterpretation of neurotoxic effects, the less protective 50% BDNF reduction was proposed as the POBEL, because severe effects were observed at this threshold in all four proposed key studies providing additional certainty in the interpretation of effects. In a weight of evidence approach, the missing BDNF information in blood from animal studies was combined with the data provided by the occupational study conducted by Zou et al. 2024.. If possible, the measurement of (D)NT effect-biomarkers in blood samples of workers should be performed, improving the knowledge of these effect-biomarkers in occupational assessments. Currently, the overall picture for this relevant effect biomarker is convincing, however data in occupational settings are limited. The inclusion of BDNF biomarkers in occupational studies is the next step needed to allow a progressive refinement of the current POBEL or to derive an OBEL in subsequent evaluations. While it should be noted that BDNF levels measured in serum reflect both the levels coming from the nervous system as well as those originated inside platelets in bone marrow cells (Boukhatem I. et al. 2024), serum BDNF levels is to date the most widely used and validated BDNF biomarker, and has been shown to predict diverse neuropsychiatric and neurodegenerative conditions (Stenz et al. 2015). The evidence also shows that biomarkers measuring blood BDNF DNA methylation at different Exons could be more specific to the central nervous system (Kundakovic et al. 2015; Stenz et al. 2015). Interestingly, among adolescent boys, Mustieles et al., measured both BDNF serum levels and blood DNA methylation levels in Exon IV of the BNDF gene, revealing that, as expected, higher DNA methylation was significantly correlated with lower BDNF serum levels (Mustieles et al. 2022), supporting that serum BDNF levels also reflect levels in the nervous system. While in general the evidence is currently higher for serum BDNF levels (at the technical, epidemiological and clinical levels), when possible, it would be advisable to incorporate additional measures such as BDNF DNA methylation and neurocognitive tests, to help understand the specific effects of exposure to a particular chemical on the human nervous system.

The Brain-derived neurotrophic factor (BDNF) is synthesized as a proBDNF within the endoplasmic reticulum of neurons and other brain cells. This proBDNF is then transported through the Golgi apparatus and stored in either dendrites or axons and then proBDNF undergoes cleavage intra or extracellularly to produce a mature, active BDNF protein. Several studies have demonstrated that BDNF can be measured in exosomes of neuronal origin, isolated from plasma using neuronal surface markers (e.g., L1CAM). This approach significantly improves CNS specificity by enriching for vesicles originating from neurons and may help separate peripheral contributions from brain-derived BDNF. While still in early stages of clinical validation, this additional methodology could allow BDNF to be reconsidered as a potential neurotoxicity marker, provided that rigorous pre-analytical and analytical controls are in place.

Justification for OBEL

Any qualitative relationship of BDNF to adverse effects of higher incidence of the adverse outcome, AO is described in the relevant AOPs (https://aopwiki.org/wiki/index.php/Aop:12 ; https://aopwiki.org/aops/13; https://aopwiki.org/aops/54) only in animal or *in vitro* studies (e.g., cell death, changes in neuronal morphology, alterations in synaptogenesis or neuronal activity etc.). Although there is a lack of

occupational studies examining Pb exposure in relation to BDNF biomarkers to be able to set an OBEL, several epidemiological studies of Pb exposure and cognitive impairment are available where association between Pb levels in blood and decreased neuronal function or cognitive impairment is established (see Table A 4.11 in A4 Appendix). Moreover, a recent study indicate links between alterations of BDNF levels and exposures of a variety of chemicals (see table 2 in Rodriguez-Carillo et al. 2024). Among modes of action, antagonism of Pb on N-methyl-D-Aspartate receptors (NMDARs) is known to be a molecular initiating event (MIE) leading to altered BDNF regulation in the brain that may result in learning and memory impairments (Sachana et al. 2018; Karri et al. 2016).

A4 Appendix:

Table A 4.10. Total serum BDNF levels found in the scientific literature from epidemiological studies with the larger sample sizes

Author	Study Design	Population	Mean/ Median BDNF level	Methodology	Conclusions
Naegelin et al. 2018	Methodological and validation study: Development of an ELISA to determine serum BDNF levels in a cohort of volunteers (n 259) and to monitor these levels in most of these volunteers (n 226) 1-year after recruitment to test BDNF stability.	259 healthy participants (178 females, 81 males) Age: 44.3 y. (range 18-70) Ethnicity: Northern Europe Blood collection between 8am and 12pm for 75% of participants. It seems both fasting and non-fasting blood samples were collected.	Recruitment Female 32.9 (SD 8.6) ng/mL Males 32.3 (SD 8.5) ng/mL Follow-up Female 33.0 (SD 8.5) ng/mL Male 33.0 (SD 8.2) ng/mL Range: 16 – 80 ng/mL No min-max or percentiles reported	Given reports on the heterogeneity of values obtained using commercially available BDNF ELISA kits (Polacchini et al., 2015), a slightly modified version of a previously described ELISA protocol was used (Kolbeck et al., 1999). Combination of BDNF monoclonal antibodies designated mAb BDNF- #1 and mAb BDNF-#9 (Kolbeck et al. 1999). Intra-assay CV of 8.8% based on a single serum sample measured on 18 different plates Intra-assay CV of 9.7% based on 253 samples tested in triplicate. The ELISA results were validated using Western Blot.	• No significant difference in mean serum BDNF levels after one year • Half of the individuals retested after 1 year had BDNF values within 10% of their levels at recruitment. 37 individuals showed >20% changes. • Power analysis indicated that a group size of 60 participants is needed to detect a 20% change in serum BDNF levels with 80% of power. • Age was weakly and positively correlated with BDNF levels.

Fonseca-Portilla et al. 2019	Aim: Examine the associations of BDNF serum levels with BMI, physical activity, and the rs6265 polymorphism among 349 Latinos aged ≥18 years enrolled in the Arizona Insulin Resistance Registry.	349 participants (233 women and 116 men) Mean age: 42y. Range 18-85 y. Ethnicity: Latino/ Mexican American 195 participants (56%) reported engaging in regular physical activity (*i.e.*, "active group")	Mean (SD) BDNF levels did not differ (P = 0.89) between men (33 ± 16 ng/mL) and women (33 ± 17 ng/mL) Mean BDNF levels did not differ significantly between the active 33 (ng/mL) and nonactive group (34 ng/mL) Higher BDNF levels were significantly associated with higher age (r = 0.11, P = 0.04) and higher 2-hr glucose level (r = 0.11, P = 0.04). No min-max or percentiles reported	Modified method using the ELISA kit from Promega Inc based on Hellweg et al. 2003. Limit of detection (LOD): 1 pg/mL. SNP genotyping for rs6265 was performed by the Assay-by-Design service (Applied Biosystems, CA.)	• Serum BDNF levels were positively predicted by age and higher glucose levels, but not the other covariates such as BMI. • Serum BDNF levels did not differ significantly between the physically active (N = 195) and nonactive group (N = 154) • Participants with the rs6265 polymorphism did not show significant differences in BDNF levels
Mora et al. 2019	Case-control study of bipolar patients vs controls	Healthy controls (n=49): 48.3 y. Euthymic bipolar (n=52): 47.5 y.	Healthy controls (n=49): 46 ng/mL Euthymic bipolar (n=52): 40 ng/mL Manic bipolar (n=32): 35 ng/mL	Sandwich-ELISA, using a commercial kit according to the manufacturer's instructions (Chemicon, Temecula, CA, USA). All samples assayed in duplicates. Intra- and inter-assay coefficients of variation were <12%.	• Bipolar patients showed significantly lower serum BDNF levels compared to healthy controls. • In bipolar patients, higher BDNF levels predicted improved cognitive abilities.

		Manic bipolar (n=32): 41.3 y. Blood samples collected between 8:00 and 9:00 am	No min-max or percentiles reported		
ADOLESCENTS					
Mustieles et al. 2022	Population Cohort study	130 adolescent boys aged 15-17 years Ethnicity: Caucasian Non-fasting blood sample	Min: 17.2 ng/mL P10: 20.3 ng/mL P25: 25.4 ng/mL P50: 31.5 ng/mL P75: 38.8 ng/mL P90: 47.4 ng/mL Max: 56.0 ng/mL	Quantikine, R&D Systems, Minneapolis, MN, USA Quality control developed within the Human Biomonitoring for Europe (HBM4EU) project. All samples were measured in duplicates in different plates. Intra- and inter-assay coefficients of variability were <5% and <15%, respectively.	• The Quantikine R&D Systems ELISA kit tested showed considerably better replicability and precision compared to the R&D Systems ELISA kit tested in the methodological study by Polacchini et al., 2015, which is not currently commercialized. • The precision of this Quantikine R&D Systems ELISA kit has been demonstrated inside the HBM4EU project (Rodríguez-Carrillo et al., 2020) • Serum BDNF was reliably measured in all samples.

					• Serum BDNF correlated with BDNF blood DANN methylation levels, but not with urinary BDNF protein levels. • BDNF DNA methylation but not serum or urinary BDNF protein levels mediated the effect of bisphenol A exposure on children's behavioral problems.
Pedersen et al. 2017 Huang et al. 2017	Cross- sectional study	447 healthy adolescents (214 females and 233 males) Mean age: 14 y. (Range 11-17 y.) Fasting blood samples Ethnicity/Country: Denmark	Female 26.9 (SD: 6.05) ng/mL Male 27 (SD: 6.34) ng/mL Neither percentiles or min-max were reported for BDNF levels.	Quantikine, R&D Systems, Minneapolis, MN, USA All samples were assessed in duplicates. Intra-assay and inter-assay variation not reported.	• Serum BDNF levels did not differ between male and female adolescents. • Serum BDNF levels were positively associated with HOMA-IR and triglycerides. • Physical activity was negatively associated with serum BDNF levels in boys but not girls.

Note: Table adapted and updated from Additional Deliverable Report AD14.6 of the HBM4EU project (Rodríguez-Carrillo et al. 2020) Report on the state of development of Task 14.3: Identification of needs for the implementation of both classical and new biomarkers of effect and decision criteria for their validation. Deliverable Report AD14.6 WP14 - Effect Biomarkers.

Table A 4.11. Epidemiological studies of association between lead levels and cognitive impairment by different ages of exposure

Adapted and updated Table 1 from Cognitive Impairment Induced by Lead Exposure during Lifespan: Mechanisms of Lead Neurotoxicity by Ortega R et al. 2021, but BDNF levels were not measured.

Population (Age Range)	Study Design and Subjects	Location	Lead (Pb) Levels	Outcomes	Reference
Children					
2–4 years	Cross-sectional study: 76 children, randomly selected from kindergarten with an average age 2.8 years (SD 1.45). Gesell Developmental Scale: *t*-test comparisons. The Achenbach Child Behaviour Checklist (CBCL): χ^2 test. Scores on each behaviour factor and the total behavioural score were analysed by the rank-sum test.	Xi'an, China	Blood: 4–246 µg/L (min max-range; differential potentiometric stripping analysis) **Control group** (Pb value < 50 µg/L and **Exposed group**, Pb value ≥ 50 µg/L	• ↓ Adaptative behavior, gross motor performance, fine motor movements, language development, and individual social behavior. • ↑ Depression, abnormal behavior, aggressions, social withdrawal, sleep problems, and destruction.	Hou et al. 2013
2–5 years	Cross-sectional study; 201 African-American children Child Behavior Checklist (CBCL): χ^2 test and *t*-test comparisons. Total Behavior Problem Score (TBPS) based upon the percentile ranking of raw scores and in the clinical range (90th percentile). Multiple regression analysis for the likelihood ratio	Baltimore, United States	Blood: 20–300 µg/L 78 children with a blood Pb level greater than or equal to 150 µg /L as the **high exposed group** (Mean 286 µg /L SD 9.3) and	• Troublesome behaviors (aggression, sleep problems, and somatic problems) • Increase in externalizing behaviors (motor activity, nonadaptive behaviors, conduct problems, inattentiveness, passivity, and hyperactivity)	Sciarillo et al. 1992

	and influential factor on the TBPS was measured from the mothers using the Center for Epidemiologic Studies Depression Scale (CES-D).		123 children as the **low exposed group** (Mean Pb blood 113 µg /L SD = 4.3)		
8–9 years	Cross-sectional study; 167 children. Wechsler Intelligence Scale for Children (WISC), Göttinger Formreproduktions-Test (GFT), Bender Gestalt-Test (German version), Benton Test, Diagnostics for Cerebral Damage Test and Wiener Reaction Device. Duisburg study t tests for correlating samples. Stolberg sample associations were tested using of stepwise multiple regression analysis.	Duisburg city and Stolberg city, Germany	Baby teeth: 1.4–38.5 µg/g Blood: 68–340 µg/L	• Deterioration of visual-motor integration • **IQ deficits (5–7 points**) • Disturbance of reaction performance • **Deficits in verbal IQ (4–7 points)**	Winneke et al. 1984
6–12 years	100 children recruited from high (n = 50) and low (n = 50) lead-polluted areas. Wechsler Intelligence Scale for Children, 3rd edition (WISC-III) Spearman's rank correlation, logistic and linear regressions to test independent predictors for impairment of cognitive function and the relationship between blood lead levels and cognitive function. Receiver operating characteristic (ROC) curve was used to calculate the best cut-off value of blood lead levels (based on the highest sensitivity with the lowest false-positive results) above which the majority of the children have cognitive dysfunction.	Cairo, Egypt	Blood: 30–280 µg/L (Median 9, interquartile range 60 µg /L) 43% of children had levels > or =100 µg/L)	• ↓ **IQ in 15 points** (mainly in verbal subsets, arithmetic and digit span) • Poor scholastic achievement • Cognitive dysfunction (flat affect, slow, and delayed responses)	Mostafa et al. 2009

Adults					
50–60 years	Cross-sectional study; 53 adults, 20 controls (Sex: 8M/12W) and 33 exposed (Sex: 13M/20W), patients from Boston Children's Hospital between 1930 and 1942 who had received a diagnosis of lead poisoning. All subjects for whom race was recorded were white and were matched for sex and age. Potential confounding factors considered. Wechsler adult intelligence scale-revised (WAIS-R), Wechsler memory scale (WMS), a test of attention and visuomotor tracking (trail making), test of verbal fluency (FAS), test of non-verbal reasoning (Raven progressive matrices), test of motor speed (finger tapping) and inventory of current mood (POMS). Wilcoxon signed ranks test, used to compare matched pairs of subjects exposed to lead and controls; and the Mann–Whitney test, used to compare entire groups. χ^2 test was used to evaluate the distribution of categorical frequencies.	Boston, United States	Blood: 60–120 µg/dL For the 12 pairs that differed in rank, T = 9.0 (p < 0 02) No blood Pb levels reported in the control group.	• Difficulties in attention and executive functioning, reasoning, and short-term memory. • Attenuation of the ability to learn new information. • Problems with attention and concentration • Reduced ability to do more than one activity at the same time. • Impaired ability to organize information or steps in a procedure. • Increased difficulty in arriving at solutions for problems. These symptoms begin at blood lead concentrations in the range 60-100 µg/dl (severity score of 1): the additional symptom in the range 90-120 µg /dl (severity score of 2) and subjects showing signs of encephalopathy (somnolence, semistupor, coma, convulsions, or projectile vomiting) begin at blood lead concentrations above 120 µg/dl, were given a score of 3.	White et al. 1993
50–70 years	Cross-sectional study; 1033 adults. A battery of 20 cognitive test results was standardized and collapsed into 7 cognitive domain scores. All 7 domain scores were standardized for	Baltimore, USA Neighborhoods	Tibia 18.8 ± 11.6 µg/g	• Affectations in language and executive function. Hierarchical mixed-effects	Glass et al. 2009

	direction so that a negative regression coefficient indicated worse performance. χ^2 test for interaction between tibia lead, Neighborhood Psychosocial Hazards Scale (NPH). NPH scale and the 7 domains of Cognitive function in the Baltimore Memory study. Multilevel regression models were used to account for the nesting of persons within neighborhoods.	were selected to offer variation by race/ethnicity and socio-economic status.	(measured via 109Cd-induced K-shell X-ray fluorescence)	regression models showed that neighborhood psychosocial hazards exacerbated the adverse associations of tibia lead in 3 of 7 cognitive domains after adjustment for age, sex, race/ethnicity, education, testing technician, and time of day (language, P = 0.039; processing speed, P = 0.067; executive functioning, P = 0.025)	
Male workers					
22 years	Case report 22-year-old man working in lead acid battery manufacturing unit	India	Blood: 128.3 µg/dL (method for Pb measurements not reported)	Lethargy, fatigue, peripheral neuropathy and weakness of forearm extensor muscles. Aggressiveness	Herman et al. 2007
39–50 years	Cross-sectional study: 47 adults of age mean 39.5 (SD 9.7) years, exposed to Pb for 11.7 ± 9 years and 53 male workers as control employed at a steel production, age 39.3 years (8.4) (mean, SD). Modified version of the Wisconsin card sorting test, the block design test, the visual recognition test, choice reaction, simple reaction, and digit symbol substitution. One-tailed *t*-test for independent samples was used for the following tasks: block design and visual recognition tests, simple reaction time, and digit symbol substitution. The results of the choice reaction were analyzed	Germany Both groups were matched on age, verbal intelligence and the same socio-economic background	Blood: Mean (SD) **Exposed:** 308 (11.2) µg/L **Control** 43.2 (20) µg/L (Atomic absorption spectrometry)	• Neurobehavioral performance poorer than control in categories such as executive functions (visual recognition), short-time memory, and visuospatial abilities.	Barth et al. 2002

	by multivariate analysis of variance. Since scores of the Wisconsin test were not distributed normally, they were analyzed by one-tailed Mann– Whitney test. Because of multiple univariate testing, Bonferroni correction was applied.	Two-tailed t-tests for independent samples showed no significant differences in age (P=0.9) and in verbal intelligence (P=0.42).			
25–67 years	Cross-sectional study: 43 workers from a lead smelter were exposed to Pb for 1–7 years compared with 45 workers from a glass factory (non-exposed to Pb). WHO neurobehavioral core test battery. Multiple linear regression of neurobehavioral function in workers and lead exposure indices. Correlation coefficients (Pearson r) between blood lead concentration and covariates Analysis of covariance for dichotomous exposure variables.	Venezuela Information on work history, medications; lifestyle exposures to neurotoxins (tobacco, alcohol, hobbies) was collected.	Blood: The range: 90–600 µg/L **Exposed group**: lead smelter: Mean: 420 µg /L (SD not reported) **Non-exposed group:** glass factory: Mean: 150 µg /L (SD not reported)	• ↓ results on Wechsler adult intelligence subtests. • ↓ Simple reaction time. • Mood indicative of depression. Anger hostility, fatigue, and confusion. 95% Confidence intervals (95% CIs) were based on normal approximations.	Maizlish et al. 1995
71 years	Case report Twins	Boston, United States	Blood: 150–1250 µg/L Patella: 119–343 µg/g Tibia: 79–189 µg/g	• Verbal/language abilities and working memory/executive function lower than average (worse in twin 1)	Weisskopf et al. 2004

			Patella and tibia bone Pb measured by K-X-ray fluorescence.	• Deficits in short-term memory function (in learning and retention of new information) • Bad manual motor control	
Female workers					
55–65 years	Cross-sectional study: 65 volunteers (31 exposed to Pb for 1.4–20.7 years and 34 healthy controls). The subjects were age and sex-matched. N-back working memory paradigm. Mean values of continuous variables were compared using the Student's *t*-test. Pearson correlation analyses between mean percentage changes of activated brain regions and working memory performance. The effects of blood lead on percent signal change by multiple regression analysis.	Korea	Blood Pb: **Exposed group:** 40.7 μg/L (80∼135) **Control group**: 20.0 μg/L (12.4∼64.7), p<0.001 Pb measured with a Zeeman background-corrected atomic absorption spectrophotometer LOQ: 0.60 μg/dL Demographic characteristic of study subjects is covered.	• Decreased working memory (in-back memory task). • Decreased activation in the dorsolateral prefrontal cortex, ventrolateral prefrontal cortex, pre-supplementary motor areas, and inferior parietal cortex	Seo et al., 2014

A5: Neurofilament light chain (NfL) in serum

The structural integrity of the dynamic network among neurons composed by axons and dendrites is realized by various scaffold proteins and the cytoskeletal protein neurofilament light chain (NfL), the smallest of the three subunits that constitute neurofilaments, is one of the important scaffolding proteins in neurons (Khalil et al. 2018). Several key events in AOPs, such as “synaptogenesis, decreased” or “neuronal network function, decreased” are directly related to the integrity and dynamics of such scaffolding proteins. While predominantly expressed in axons of the central and peripheral nervous system (Khalil et al. 2018) there is also evidence that synaptic plasticity is supported by NfL (Gafson et al. 2020). In development neurotoxicity cytoskeletal proteins like microtubule (MT), neurofilament heavy chain (NfH), neurofilament light chain (NfL), and neurofilament medium chain (NfM) have been considers as targets of oxon analogues form by organophosphorus ester (OP) insecticides (Flaskos, 2014). However, to the best of our knowledge, DNT cohort studies have never used NfL as an effect biomarker. After axonal/neuronal damage, NfL is released into the cerebrospinal fluid (CSF) or the endoneurial fluid and first approaches using NfL as a biomarker of neuronal damage analysed this protein usually in CSF. One hallmark study (Rosengren et al. 1996) showed that patients with amyotrophic lateral sclerosis (ALS) and Alzheimer’s disease (AD) had significantly increase NfL concentrations in CSF when compared to controls. Especially in ALS patients these levels were markedly elevated. Since the 1990s technical development in protein analysis made it possible to reliably measure NfL in blood/serum (sNfL). Nowadays, the fourth-generation SiMoA assay is 126-fold and 25-fold more sensitive than ELISA and ECL assays (Kuhle et al. 2016) that have previously used to analyse NfL in serum. Most of the studies that are relevant in the context of this paper are based on this technology or comparable techniques with excellent sensitivity (Analytical LLOQ: 0.640 pg/mL; Functional LLOQ in serum and plasma: 2.56 pg/mL). Compared to this excellent sensitivity, the specificity of this biomarker of effects is weak. Various neurological disease such as Multiple sclerosis (MS), Dementia, stroke, Mild traumatic brain injury (TBI) or the aforementioned disease AD and ALS are known to affect sNfL, but such patients are usually excluded of epidemiological studies addressing neurotoxicity in the working environment. However, various factors (e.g., age, smoking, BMI, etc.) are known to affect sNfL (see **Figure A 5.7**) and for aging, an ~2.2% increase in serum NfL in each year has been reported (Khalil et al. 2018).

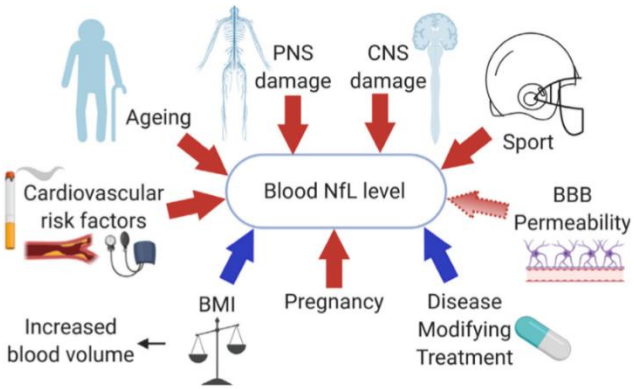


Figure A 5.7 Physiological and pathological factors increasing or decreasing the blood levels of NfL (Barro et al. 2020)

Some of the most relevant factors namely age and BMI are integrated in the NfL app (Benkert et al. 2022) that calculates z-scores from serum levels based on a control population (Web application: http://shiny.dkfbasel.ch/baselnflreference). Additional normative data has been described in two studies (Beltran, 2024; Bornhorst et al. 2022). Epidemiological studies assessing the neurotoxicity of chemicals in

the workplace y the help of sNfL alone or in combination with neurobehavioral tests are to the best of our knowledge not available yet. Nevertheless, the usefulness of this biomarker of effect could be shown in the context of paclitaxel-induced peripheral neurotoxicity (PIPN) in patients and animal models. Paclitaxel is a cytostatic anti-cancer drug approved in the United States for the treatment of breast, pancreatic, ovarian, Kaposi's sarcoma and non-small-cell lung cancers, (Balayssac et al. 2023; Velasco et al. 2023). Paclitaxel binds to ß-tubulin and thereby inhibits the degradation of spindle fibers, which are made up of it. This inhibits mitosis in the G2 and M phases, blocks cell division and triggers programmed cell death. In cancer patients, the serum NfL levels were significantly correlated with the neurophysiological phenotype of PIPN that typically manifests as a predominantly sensory axonopathy causing pain and nerve conduction (NC) abnormalities (Velasco et al. 2023). The evidence from clinical studies assessing CNS neurotoxicity of (neo)adjuvant chemotherapy for breast cancer is less striking but provides some proof-of-concept support as an CNS damage biomarker (Schroyen et al. 2023). sNfL has also been used as predictive biomarker of treatment effects in MS patients as well as for treatment evaluation (Benkert et al. 2022). MS patients treated with high efficacy monoclonal antibody therapies returned to sNfL z-scores of zero within one year of treatment. In a recently published study using the large NHANES cohort in the USA, high serum NfL levels (i.e., (≥ 19.0 pg/mL) and low cognitive function were significantly correlated (Gao et al., 2023). Finally, one exploratory study showed a decreasing cognitive performance with increasing serum NfL (Beste et al. 2019). Another recent NHANES data analysis of serum samples from 2071 adults reported mean serum NfL levels of 7.99 pg/mL (95% confidence interval [CI]=15.43–20.17) and 15.78 pg/mL (95% CI=13.00–18.55) for males and females, respectively, after controlling for age (Beltran 2024). The overall mean serum NfL concentration was 16.76 pg/mL (95% CI=14.45–19.07 pg/mL).

According to the current scientific evidence briefly described above and the plausibility of the underlying neurobiological mechanisms, all requirements for using sNfL as biomarker of neuronal damages seem to be fulfilled.

Justification for TOBEL:

In serum NfL can be easily and reliably detected in healthy and diseased subjects using fourth generation neurofilament assays (e.g. SiMoA) with a LOQ of 0.02 ng/L NF-L. In addition to serum, NfL can be measured in plasma and whole blood but most of the reference values are related to measurements obtained in serum.

Justification for ROBEL:

By using the Web application (http://shiny.dkfbasel.ch/baselnflreference) raw sNfL concentrations measured with suitable methods can be transformed into age- and BMI-adjusted z-scores. Equations are available and implemented in Excel or R. These values are normally distributed and cut-offs representation percentiles of the population could be used as ROBEL (e.g. z-score > 1.96). Such a z-score can be interpreted as being among the 5% of a population with the highest measured sNfL concertation. Several approaches and publications are available providing normative data (e.g. z-scores) that could be used as reference level (Benkert P. et al. 2022)

Justification for POBEL:

The German BAT value (150 µg/L) for lead in blood is based on NOEAL of 180 µg/L as derived from epidemiological studies of lead exposed workers (e.g. Schwartz et al. 2001). Currently, a biomonitoring study is conducted by Prof. Goen (Erlangen, Germany) measuring both blood lead and sNfL simultaneously. As soon as this data will be available, the raw sNfL data will be normalized to z-scores and associations with the blood lead levels will be analysed using different statistical methods. Finally, a cut-off value separating workers below and above the BAT value might serve as new POBEL.

As intermediate POBEL and based on (Benkert, P. et al. 2022) NfL levels above age specific 95% can already indicate a reduced structural integrity of the brain (see **Table A 5.12**).

Table A 5.12. Age dependent POBEL for NfL

Age classes	NfL (95% percentile [ng/L])
41-50	12-15
51-60	15-20
61-70	20-26
71-80	26-32

These age-specific findings are supported by a recent study (Rodero-Romero et al. 2024) and are also in agreement with Gao et al. 2023 showing that values in elderly (aged 60 or older) with sNfL levels ≥ 19 ng/L showed reduced performance in a cognitive test (i.e. digit symbol substitution test, DSST).

Justification for OBEL

Several quantitative relationships of sNFL concentrations with neurological disease severity (e.g. MS, TBI, anti-cancer drugs) and cognitive performance are available, but a derivation of an OBEL was not yet performed. More research and data including traditional neurobehavioral endpoints of neurotoxicity (e.g. motor and cognitive performance tests) are needed to come to a justified OBEL.

More information is available in the **Annex B6**.

A6: Neurogranin (NGRN, Ng) assessment in serum

Neurogranin (NGRN, Ng) is a neural-specific postsynaptic protein that is highly enriched in dendrites of excitatory neurons of the cerebral cortex, hippocampus, and striatum (Pak et al. 2000). When NMDA receptors in dendritic spines are activated, calcium ions enter the postsynaptic neuron and release Ng-bound calmodulin (CaM). Free CaM then activates a network of calcium-dependent kinases and downstream signaling pathways (Bernal 2009). These pathways are involved in synaptic plasticity and long-term potentiation, a process that strengthens the synapses between neurons, which represents the cellular basis of learning and memory (Dong et al. 2015).

Patients with Alzheimer's disease (AD) have markedly reduced levels of Ng in the hippocampus and frontal cortex, indicating synaptic loss, which occurs early in the disease course and leads to reduced cognitive function. The occurrence and progression of AD have been shown to be closely related with Ng levels in cerebrospinal fluid (CSF) and blood (Liu et al. 2020). Meta analysis of epidemiology data reported a significantly elevated CSF Ng in AD patients and the people with mild cognitive impairment (MCI) compared to healthy control subjects. Furthermore, the CSF Ng levels were higher in AD patients than the MCI patients (Liu et al. 2020). It should be noted that Ng may reflect a general pathophysiological process of synaptic degeneration implicated in several neurological disorders and not only specifically to AD (Agnello et al. 2021). Other studies indicated that the raised Ng levels in CSF are correlated with structural measures of hippocampal atrophy as measured by functional PET-MRI imaging systems (Portelius et al. 2015). Contrary to CSF Ng, the concentrations of Ng in blood (plasma exosomes) of AD patients and MCI patients were lower than that of healthy controls (Liu et al. 2020). High serum Ng has been detected in acute conditions where brain tissue is damaged, such as traumatic brain injury, stroke, brain infection, epilepsy, and CO poisoning (see **Table A 6.13**). Overall, the available clinical evidence suggests that Ng in CSF and plasma/serum can be used as a useful biomarker of synaptic dysfunction and cognitive impairment to assess the process of neurodegeneration (Liu et al. 2020). However, the clinical value of Ng is not entirely clear yet. As for DNT, few experimental studies examined changes in Ng expression following developmental exposure to chemicals. A sustained reduction in the number of Ng-immunoreactive cells (Ng^+=: Neurogranin positive (Ng^+) cells are neurons that are specifically identified by the presence of Ng protein in immunocytochemistry assays) was observed in the rat hippocampus after developmental exposure to propylthiouracil, glycidol and ethanol (Takahashi et al. 2023). Prenatal stress also reduced Ng expression in rat brain (Sivasangari et al. 2020).

Justification for TOBEL:

Since synaptic dysfunction is associated with cognitive decline and neurodegeneration, biomarkers reflecting the integrity and plasticity of synapses, such as Ng, may be useful for the early diagnosis and prognosis of chemically-induced neurodegeneration in occupational settings, where exposure to metal elements, organic solvents, gases, or pesticides poses a risk.

The concentration of Ng has been measured in various biological fluids using a variety of analytical methods (see **Table A 6.13**). Immunoassays, in particular ELISA and electrochemiluminescence are the most commonly used, while other more precise methods (e.g., hybrid immunoaffinity-mass spectrometry) have been used less frequently. CSF and plasma or serum are the most commonly used biological samples, and their Ng concentration levels are expressed using different units (ng/mL, ng/dL, pg/mL). A. LOD around 0.05 ng/L has been reported for serum Ng using commercial immunoassays (see **Table A 6.13**). Methodological issues should be considered for comparative purposes. CSF has the advantage of more accurately reflecting changes in brain tissue; however, it is an invasive sample that can only be obtained from hospitalized patients. Therefore, more readily available, and less invasive samples, such as serum or plasma, are being used more frequently. Since Ng is expressed in the lung, spleen, bone marrow, and platelets, these extraneural sources may contribute to blood Ng concentration (Xiang et al. 2020), thus resulting in lower specificity and confounding results. This limitation can be overcome by using plasma-

derived neuronal exosomes that can be enriched from blood samples and better reflect brain pathogenic processes (Camporesi et al. 2020). These exosomes, secreted by neurons and released into the blood (aka neuronal-derived extracellular vesicles, NDEV), offer greater specificity for brain tissue and have been used in different studies for measuring Ng in plasma as shown in **Table A 6.13**. Furthermore, foetal NDEV have been isolated from maternal plasma as a non-invasive platform for testing impaired foetal neurodevelopment in early pregnancy (Goetzl et al. 2019).

Justification for ROBEL:

The limited available evidence indicates that Ng levels in CSF increase with age and are higher in females than in males (Milà-Alomà et al. 2021). However, these observations have been made in patients and it is unclear whether these differences will also be found in the general population.

In the adult control population of the studies shown in **Table A 6.13**, serum Ng levels ranged from 0 to 200 ng/mL. The largest study (n=328) reported a mean value of 14.9 ± 30.3 ng/mL (Peacock et al. 2017). For studies measuring plasma NDEV, levels ranged from 0.28 to 0.63 ng/NDEV in 1 mL plasma. The largest study (n=160) reported a mean value of 0.63 ± 0. 17 ng/mL (Jia et al. 2021).

However, laboratory test results have no clinical value in isolation and, consequently, reference interval and decisional cut-off value are mandatory for appropriately interpreting the laboratory data. Agnello et al. (2021) established the reference interval of Ng in CSF from a population of 100 controls and individuals with non-neurodegenerative neurological diseases. The lower and upper limits of the reference interval of Ng in CSF, calculated by the robust method with bootstrapped 90% confidence interval (CI), were 2.9 (0.1–10.8) and 679 (595–779) pg/mL, respectively. The diagnostic performance analysis of Ng for AD indicated that the best-calculated cut-off of Ng for diagnosing AD was 319 pg/mL. At this cut off, sensitivity, specificity, positive predictive value, and negative predictive values were 0.73, 0.73, 0.46 and 0.89, respectively (Agnello et al. 2021). A moderate positive correlation between CSF and serum Ng levels has been observed ($r = 0.615$, $P < 0.05$) (Canturk et al. 2022).

Justification for POBEL:

It is becoming increasing clear that occupational and environmental exposures to neurotoxic chemicals may influence the progression or age at onset of neurodegenerative diseases. Changes in Ng levels may reflect ongoing synaptic degeneration linked to cognitive impairment and may serve as biomarker of synaptic integrity in chronic neurodegenerative disorders (Blennow and Zetterberg 2018, Krishna et al. 2023).

A direct correlation between Pb and Ng levels in blood from workers would allow a potential neurotoxic risk to be identified and bridge occupational exposure with assessment of potential cognitive decline using Ng as effect biomarker (the updated Occupational biomonitoring Level (i.e., German BAT value) of 150 µg Pb/L in whole blood could serve as a NT anchor for adverse effects). However, this value is not protective for offspring of female lead-exposed workers at childbearing age (EU-Commission report 2023, see: https://www.europarl.europa.eu/doceo/document/A-9-2023-0263_EN.pdf). Unfortunately, occupational epidemiological studies published to date have not measured Ng levels. The incorporation of Ng as a biomarker of effect in these studies would significantly strengthen our understanding of how exposure to occupational chemicals (including metal elements, organic solvents, and pesticides) impact neurotoxicity.

Two animal studies have reported changes in Ng expression after metal exposure. Prenatal methylmercury induced downregulation of Ng protein expression in the offspring rat (Jacob and Sumathi 2019). Perinatal cadmium exposure reduced serum thyroid hormone (TH) levels in mice pups, and as a result, decreased the Ng gene expression, which is regulated by TH through TH receptors (Ishitobi et al. 2007).

Justification for OBEL

To the best of our knowledge, Ng has not been identified as a key event, and there is no documented quantitative relationship between Ng levels and apical adverse outcomes in relevant AOPs within the AOP Wiki (https://aopwiki.org/). However, Ng is mentioned in AOP 42: I*nhibition of Thyroperoxidase and Subsequent Adverse Neurodevelopmental Outcomes in Mammals*. In that AOP, decreased foetal expression of Ng in the rat hippocampus and cortex serves as an example of empirical evidence for KER 1387 "*T4 in serum, Decreased leads to Hippocampal gene expression, Altered*" (https://aopwiki.org/relationships/1387). Ng is also mentioned in a reference supporting KE 756 "*Hippocampal gene expression, Altered*" (https://aopwiki.org/events/756).

More information about this effect biomarker is available in **Annex B7**.

Overall assessment of the analytical performance of the methodology used to assess serum Ng levels and other sources of variability.

The methodology most commonly used for measuring Ng in the studies shown in **Table A 6.13** was ELISA; however, most studies did not provide information on the analytical performance of the method used. Authors were more interested in diagnostic performance of the tests than in the analytical performance. The reliability of the results depends on whether the analytical method has been validated or at least the study refers to a previously validated method. Criteria for validation include description of the nature of antibodies used, LOD, LOQs (lower and upper), inter-assay and inter-plate (i.e., inter- and intra-assay, respectively) coefficients of variation, recovery, calibration curve (range and linearity) and whether the analysis was performed by duplicate or triplicate. Only two studies shown in **Table A 6.13** reported most of these criteria (Peacock et al. 2017 and Yang et al. 2015); the rest reported only some of these criteria or even none, which limits their validity. For results to be valid they must be comparable with results of the same ELISA or ECLIA (electrochemiluminescence immunoassay) performed in different laboratories (inter-laboratory reproducibility). Using hybrid immunoaffinity-mass spectrometry (HI-MS), several plasma Ng peptides were reliably detected and quantified (Kvartsberg et al. 2015). Furthermore, since the studies reviewed addressed diverse neurological diseases (such as neurodegenerative, traumatic or intoxication diseases), the concurrent controls varied by age, sex, race, country and other factors, which also contributes to the lack of comparability of results and make it difficult to identify reference values applicable to the general population.

A6 Appendix:

Table A 6.13. Levels of Neurogranin (Ng) in human plasma/serum samples in studies drawn from the scientific literature.

First Author*	Study design	Country	Age (years) / (n; % ♀)	Ng levels		Disease	Methodology (LOD)	Method validation	Units
				Control	Patients				
Abner et al. 2016	Longitudinal cohort (volunteers)	USA	1st sample*: 69.3 ± 5.6 y (n=18; 55.6% ♀) 2nd sample*: 77.6 ± 6.1 (n=20; 50.0% ♀)	1st sample*: 0.285 ± 0.077	2nd sample*: 0.225 ± 0.061	Older cognitively intact subjects (two time-points)	Human-specific ELISA	Not reported	Plasma NDEV (ng/mL)
Jia et al. 2021	Longitudinal dataset	China	**Controls**: 60 ± 4 y (n=160, 48.9% ♀) **Preclinical AD**: 60 ± 4 (n=160, 51.1% ♀)	0.628 ± 0.166	0.254 ± 0.069 #	Preclinical AD	ELISA kits	Not reported Calibration curve: 0.039-10 ng/mL	Plasma NDEV (ng/mL)
Winston et al. 2016	Cross-sectional	USA	**MCI** : 68.70 ± 7.76 y (55-79) y (n=20, 53.8% ♀) **ADC**: 75.35 ± 6.82 y (63-87) y (n=20, 81.8% ♀)	0.357 ± 0.026	MCI: 0.254 ± 0.028# ADC: 0.056 ± 0.013#	AD, MCI, ADC**	ELISA	Not reported	Plasma NDEV (ng/mL)

			Control: not reported		AD: 0.050 ± 0.011#				
Alvarez et al. 2022	Randomized clinical trial	Spain	**Control** : 74.35 ± 5.76 y (n= 20, 75% ♀) **AD**: 74.87 ± 7.55 y (n= 116, 80.2% ♀)	0.499 (CD81-adjusted)	0.288# (CD81-adjusted)	AD	Human-specific ELISA	Not reported	Plasma NDEV ng/mL (CD81-adjusted levels)
Álvarez et al. 2022	(see above, same study)	(see above, same study)	(see above, same study)	5.78 ± 0.88	5.20 ± 0.93#	AD	human-specific ELISA		Plasma NDEV per nl (natural log values)
De Vos et al. 2015	Cross-sectional	Belgium	**Control:** 48 years (42–62) (n=29, 62.1% ♀) **MCI**: 78 years (73–82) (n=20; 65.0% ♀) **AD**: 77 years (71–84)	Median: 1.29 IQR: 0.15–3.78	MCI: Median 0.88 IQR: 0.03–2.91 AD: Median: 0.14# IQR: 0.05–1.01	AD, MCI	In-house ELISA Quantification not possible in all samples (particularly in controls).	Lower LOQ: 0.003 ng/ml. mAb Ng7. Specificity assessed. Calibrator curve (synthetic full-length Ng, 4-fold dilutions from 2 ng/mL).	Plasma (ng/mL)

			(n=20; 55% ♀)					By duplicate	
Kvartsberg et al. 2015	Cross-sectional	Germany	**Control**: 54 ± 14 y (41–63) (n=20; 60.0% ♀) **AD**: 76 ± 8 (76 (71–78) (n=25; 56.0% ♀)	Median: 47.5 IQR: 21.9–90.3	Median: 36.5 IQR: 25.3–57.7	AD	In-house immunoassay on the Meso Scale Discovery (MSD) platform	mAb Ng7. Calibration curve (0.031 to 23 ng/mL). Blanks (zero concentration)	Plasma (ng/mL)
Kvartsberg et al. 2015	(see above, same study)	(see above, same study)	(see above, same study)	Median: 21698 IQR: 13.4–107.0	Median: 25.6 IQR: 17.3–53.9	AD	Hybrid immunoaffinity-mass spectrometry (HI-MS)	Cross-reference to the HI-MS analytical method. Non-specific binding checked.	Plasma (ng/mL)
Shang et al. 2022	Case-Control (forensic autopsies)	China	**Controls** : 51.0 ± 26.2 y (n=30: 26.7% ♀) **Forensic TBI** : 47.6 ± 26.2 y (10–87) (n=56; 30.4% ♀)	0.033 ± 0.020	0.139 ± 0.062 #	TBI	Human ELISA commercial kit	Assay range: 9.38–0.6 ng/mL Calibration curve. By duplicate	Serum (ng/mL)
Yang et al. 2015	Case-Control	USA	**Controls** : 54 y (47–62) (n=150; 52.7% ♀)	Median: 0.02 IQR: 0.05–0.07	Median: 0.18 # IQR: 0.05–0.64	TBI	In-house ECLIA (electrochemiluminescence sandwich immunoassay)	Validation study: Antibodies description.	Serum (ng/mL)

	(hospital-based)		**TBI** : 47 y (30–56) (n=76; 38.2% ♀)				on Meso Scale Discovery (MSD) platform	LOD: 0.055 LLOQ: 0.200 Interassay CVs ≤ 10.7% Calibrators: 40.0–0.055. Average recovery: 99.9% (97.2–102%)	
Peacock et al. 2017	Case-Control (hospital-based)	USA	**Controls** : 38.96 ± 13.21 y (n=328; 56.7% ♀) **mild TBI**: 42.70 ± 17.20 y (n=334; 34.4% ♀)	14.93 ± 30.27	12.39 ± 28.08	Mild TBI	ECLIA (electrochemiluminescence immunoassay) on Meso Scale Discovery (MSD) platform	Validated. Antibodies description. Acceptance criteria: CV <10%, recovery 80–120%, linearity <99%. LOD: 0.041	Serum (ng/mL)
Çevik et al. 2019	Case-Control (hospital-based)	Turkey	**mild TBI**: 24 ± 22 y (5–65) (n=48; 20.8% ♀) CT scan – (n=24) CT scan + (n=24)	2.95 ± 2.38 Median 1.8 Range: 1.3–9.52	5.79 ± 4.14# Median 3.97 Range: 2.06–13.04	TBI (CT scan+)	ELISA commercial kit	Not fully reported. LOD: 0.051 Range 0.1–30 CV <10%	Serum (ng/mL)

De Vos et al. 2017	Cohort of AIS patients from a prospective study	Belgium	**Acute ischemic stroke (AIS)** 71 ±14 years (n=50, 46.0% ♀)	No control group	Median: 1.35 IQR: 0.82–2.48	Acute ischemic stroke (AIS) (at admission)	In-house ELISA	Referred to other paper: mAb, Ng fragments truncated at P75 were quantified. Interassay CV <5% and inter-plate CV < 14%. By duplicate	Plasma (ng/mL)
Kuşdoğan et al. 2023	Case-Control (hospital-based)	Turkey	**Controls** : 67.24 ± 9.87 y (n=55; 58.2% ♀) **Cases (AIS)** 69.15 ± 11.86 y (n=86; 51.2% ♀)	Median: 121.26 IQR: 90.35	Median: 160.00 # IQR: 75.93	Acute ischemic stroke (AIS)	ELISA commercial kit	Analytical performance not reported	Serum (ng/mL)
Canturk et al. 2022	Case-Control (hospital-based)	Turkey	**Controls**: 29 ± 18 (18–49) (n=15; 53.3% ♀) **Cases**: 37 ± 28 y (20–80) (n=15; 26.7% ♀)	198.6 ± 51.7	429.2 ± 104.3#	CNS infections	ELISA commercial kit	Analytical performance not reported	Serum (ng/mL)

Kalkan et al. 2022	Case-Control (hospital-based)	Turkey	**Controls**: 33# (range 18–67) (n=28; 47.2 % ♀) **Cases**: 29# (range 18–60) (n=49; 61.2% ♀)	Median: 0.98 Range: 0.74–2.82	Median: 1.84# Range: 1.10–11.73	Epileptic seizures	ELISA commercial kit	Analytical performance not reported	Serum (ng/mL)
Yeşilyurt et al. 2021	Case-Control (hospital-based)	Turkey	**Controls** : 48.91 ± 18.63 y (n=32; 56.3% ♀) **Cases** : 51.61 ± 17.65 y (n=36; 52.8% ♀)	0.22 ± 0.10 Range: 0.04–0.38	0.31 ± 0.16 Range: 0.14–0.98	CO poisoning	ELISA commercial kit	Analytical performance not reported	Serum (ng/mL)

*Full references are cited in the References A6

Median values (with interquartile range, where appropriate) or mean ± standard deviation (# p<0.05)

AD: Alzheimer's disease; MCI: mild cognitive impairment; ADC: MCI converting to AD

TBI: traumatic brain injury.

ECLIA (electrochemiluminescence immunoassay) on the MSD (Meso Scale Discovery) platform

NDEV: Neuronal-derived extracellular vesicles enriched for neuronal origin

*Two samples collected at 3- to 11-years intervals in older cognitively intact subjects (plasma NDEV levels)

IQR (Interquartile Range), mAb (Monoclonal Antibody), LOD (Limit of Detection), LOQ (Limit of Quantification), LLOQ (Lower Limit of Quantification), CO (Carbon Monoxide), CV (Coefficient of Variation)

A7: Reproduction toxicity assessment in occupational biomonitoring. Male testosterone level assay

Name of the effect biomarker: total serum testosterone level (TT)

Positive control in dose relationships: DEHP

Summary

The use of total (bound and unbound) serum testosterone (TT) levels as an effect biomarker for assessing reproductive toxicity in occupational biomonitoring is supported by various studies. Testosterone, the primary male hormone, is crucial for sex differentiation, male characteristics, spermatogenesis, and fertility. It is mostly bound to proteins like sex hormone-binding globulin (SHBG) in the blood, with only a small free fraction being bioavailable. TT levels can indicate reproductive health problems (in men), with low levels linked to adverse outcomes such as decreased fertility and abnormal sexual differentiation. TT is typically measured in serum using methods like immunoassays and liquid chromatography tandem mass spectrometry. Normal TT levels vary by age, with levels generally declining in older men. The variability in TT levels can be influenced by factors such as circadian rhythms, body weight, stress, and health conditions, making it essential to standardize sample collection, which is typically done in the morning. The biological relevance of TT levels is based on its role in several Adverse Outcome Pathways (AOPs) related to male reproductive health. These AOPs provide a framework for understanding how environmental and occupational exposures can disrupt endocrine function, leading to reduced TT levels and potential reproductive harm. Although the use of TT levels in biomonitoring has been shown for exposure to certain chemicals, such as phthalates, which are correlated with lower TT levels, the strength of the association between TT levels and adverse reproductive outcomes varied across studies. TT levels could be used to develop provisional occupational biomonitoring levels, though access to comprehensive datasets, such as those for sperm quality from WHO, is needed for further analysis of outcomes.

General introduction on reproductive toxicity

Testosterone is the primary male hormone responsible for regulating sex differentiation, producing male sex characteristics, spermatogenesis, and fertility. It is a potentially ideal biomarker for a tiered approach (OBEL, POBEL, etc.), as presented below.

Description of TT

Testosterone is released into the general circulation and transported in serum. In men, out of total serum testosterone (TT), 44 to 65% are bound to sex hormone-binding globulin (SHBG), 33 to 54% are bound to albumin and <2% are in unbound or “free” form. In women, 66 to 78% of testosterone is bound to SHBG, 20 to 32% bound to albumin and approximately 1% is in free form (Emadi-Konjin et al. 2003, Wheeler et al. 1995). While the free and albumin-bound fractions are considered to be biologically active, the SHBG bound testosterone acts as an active reservoir to buffer testosterone fluctuations in healthy individuals (Handelsman 2020). The testosterone plays an important role in the development of the male reproductive system and is necessary for normal male fertility. Low testosterone levels are associated with adverse conditions (e.g., abnormal sexual differentiation, decreased fertility) in men (Radke et al. 2018). In females, testosterone acts as an estrogen precursor. In both sexes, testosterone exerts anabolic effects and influences behavior (Handelsman 2020).

Measurement of serum testosterone levels

Testosterone is generally measured in serum as TT levels, which is the sum of the concentration of protein-bound (both SHBG and albumin) testosterone and free testosterone (Guzelce et al. 2014). If the SHBG

levels are not normal, the TT may not accurately represent the bioavailable testosterone. Thus, an SHBG measurement is recommended along with TT test (Sizar et al. 2023).

Analytical method

The commonly used analytical methods for measuring TT levels are immunoassays, liquid chromatography tandem mass spectrometry and electrochemical methods.

LOQ for serum testosterone assays is around 20 - 50 ng/L (Matsumuto et al. 2004, Sun et al. 2020). This is far below mean TT levels in 40-70 years old men, which is in the range of 4460 to 4610 ng/L (see **Table A 7.14**) (Travison et al. 2017).

Total testosterone levels in the general population

TT level is age dependent but well-investigated; see (Mohr et al. 2004, Travison et al. 2017, Clark et al. 2018).

Table A 7.14. Total serum testosterone levels in men in the general population of US and Europe (N=9054) (data from Travison et al. 2017)

Age [years]	TT level [ng/dL]		
	5%	Mean	95%
19-39	273	507	834
40-49	243	461	813
50-59	222	446	812
60-69	221	446	812
70-79	220	446	812
80-99	203	446	812

Serum TT levels in females are much lower than in males; the median serum TT level in healthy females 15 to >18 yrs of age is approximately 26 ng/dL (Clark et al. 2018)

Sampling strategy

Blood sampling should be in the morning, usually within 3 hrs of waking up, preferably in a fasting state. TT can vary as much as 10 to 15% in an individual. This variability is considered acceptable. Men who consistently work night shifts can have their testosterone levels determined in blood samples drawn at their normal wake-up times. Men who inconsistently work night shifts should have their testosterone levels measured in blood drawn between 7:00 and 9:00 am on a day-off from work.

Confounders

The TT level in serum can be confounded by various factors, including body weight, occupational stress, health effects (especially diabetics, hypothyroidism, etc.) circadian rhythms, smoking, alcohol, BMI, and use of drugs, especially drugs containing steroids.

Biological plausibility and AOP relevance

Testosterone levels are mentioned in several AOPs with plausible MoA with key events related to TT levels (Radke et al 2018, Baken et al. 2019, Arzuaga et al. 2020).

AOP# 18: PPARα activation in utero leading to impaired fertility in males

AOP#19: Androgen receptor antagonism leading to adverse effects in the male foetus (mammals)

AOP#124: HMG-CoA reductase inhibition leading to decreased fertility

AOP#288: Inhibition of 17α-hydrolase/C 10,20-lyase (Cyp17A1) activity leads to birth reproductive defects (cryptorchidism) in male (mammals)

(AOP#346: Aromatase inhibition leads to male-biased sex ratio via impacts on gonad differentiation

AOP#376: Androgen receptor agonism leading to male-biased sex ratio)

These AOPs are for prenatal testosterone only. AOP for adult testosterone does not currently exist.

Application in biomonitoring studies (occupational)

Testosterone measurements have been applied in several biomonitoring studies e.g. (Chang et al. 2015, Baken et al. 2019, Henrotin et al. 2020, Zeng et al. 2022, Zheng et al. 2022, Bracci et al. 2023). Higher exposure to several phthalates (DEHP, DINP, DIBP) is associated with lower serum total testosterone levels in exposed humans (Radke et al. 2018)

Table A 7.15. The tiered approach for serum total testosterone (TT) levels. The reference population used for the ROBEL was predominantly from white male in the U.S. and Europe (Travison et al. 2017). The TOBEL refers to the LOQ for total testosterone analyzed using liquid chromatography tandem mass spectrometry method (Sun et al 2020)

Human Health	Level	Meaning if below	Total Testosterone (TT) in serum
OBEL=Occupational Biomonitoring Effect Level	Refined	Health risk indicated	
POBEL	Provisional	Health risk may be indicated	Recommended to be derived with WHO datasets
ROBEL	Reference	Exposure to TT lowering substances is indicated if value is below reference level	Serum TT level is age depending (5%/Mean/95%) Age 19-39: 273/507/834 ng/dL Age 40-49: 243/461/813 ng/dL Age 50-59: 222/446/812 ng/dL Age 60-69: 221/446/812 ng/dL Age 70-79: 220/446/812 ng/dL
TOBEL	Technical	Exposure indicated (e.g.<LOQ or <LOD)	LOQ for TT is around 10 ng/L

Free Androgen Index (FAI)

The Free Androgen Index (FAI) is a calculated ratio used to estimate the amount of biologically active testosterone in the body. FAI is calculated by dividing TT by SHBG and multiplying this value by 100. Higher FAI indicates a higher proportion of biologically active testosterone (low SHBG levels or high TT levels) and lower FAI indicates a lower proportion of biologically active testosterone (high SHBG levels or low TT levels). A limitation is that FAI is an estimate of active testosterone. A benefit using FAI is that it can be used in both men and women, particularly in assessing androgen levels in women.

More information about this effect biomarker is available in Annex B7.

Discussion:

Quantifying the relationship between TT levels and adverse effects

Human TT level is a sensitive and well investigated clinical biomarker of effect. This effect biomarker can be used to identify in collective assessments, occupational exposures that reduce testosterone levels. Relevant AOPs for TT may be used to interpret exposure to endocrine disrupting chemicals that significantly lower the male workers' TT levels. The TT levels may thus, be used as an indication of possible adverse health endpoints.

Link between TT levels and adverse effects

There was a medium strength of association between TT levels and adverse effects in a systematic review for phthalate exposures. Radke et al 2018 classify the evidence as "robust" for effects of DEHP and DBP on reproductive endpoints, based primarily on anogenital distance (AGD), testosterone and semen parameters following DEHP exposures, and on semen parameters and time to pregnancy following DBP exposures. Indeed, more than half of the included studies report an association of DEHP and DBP exposure with AGD, serum testosterone, and sperm concentration and motility. However, up to half of the studies reported no associations. In human studies, the TT levels are measured but the exposures to chemicals are often not completely characterized. It is therefore unclear whether any of these endpoints (serum testosterone, semen parameters) is sufficiently predictive to be useful as an effect biomarker for endocrine disrupting chemicals with short urinary elimination half-lives where sampling collection time was not standardized across studies.

We assume there is a correlation between TT levels and adverse sperm quality parameters as defined by WHO, which could be used to derive a probabilistic Provisional Occupational Biomonitoring Level. During this OECD project we were not able to get access to WHO datasets that would allow a POBEL derivation. We therefore recommend that such an analysis should be conducted after having published these guiding principles.

Forms of testosterone

Different forms of testosterone are present in the blood. The majority of testosterone is bound to SHBG (60-70%), some bound to albumin (usually weak binding) and some are found in free form (1-2%). Both free and albumin-bound testosterone are bioavailable. Testosterone bound to SHBG is essentially inactive and acts as a reservoir. A high SHBG level means it is likely that less free testosterone is available to the tissues than is indicated by the total testosterone test. The primary factor that is associated with increased SHBG is aging. For example, the TT levels do not change from age 50 to 99 year (see Tables **A 7.14 and A 7.15**), but the free and bioavailable testosterone levels are in general, lower in elderly men due to increased SHBG levels with age (Emadi-Konjin et al. 2003). The other key factors that increase SHBG include certain health conditions (cirrhosis and hepatitis) and use of certain anticonvulsants (Bhasin et al. 2018). When the SHBG level is low, more of the TT is expected to be bioavailable. Obesity (body mass index >30 kg/m^2), diabetes, nephrotic syndrome, and use of glucocorticoids, some progestins, and

androgenic steroids are common conditions that decrease SHBG levels (Bhasin et al. 2018). While multiple studies have shown a clear association between obesity and reduction in serum TT, obese individuals may maintain free TT within the reference range (Fui et al. 2014). The main cause of low TT in obese individuals is the low amount of SHBG-bound TT due to less SHBG in obesity (Fui et al. 2014).

References A1-A7:

A1: In vitro Estrogen Receptor alpha (ERα) transactivation and protection levels

- Amir S, Shah STA, Mamoulakis C, Docea AO, Kalantzi OI, Zachariou A, Calina D, Carvalho F, Sofikitis N, Makrigiannakis A, Tsatsakis A. Endocrine Disruptors Acting on Estrogen and Androgen Pathways Cause Reproductive Disorders through Multiple Mechanisms: A Review. Int J Environ Res Public Health. 2021 Feb 4;18(4):1464. doi: 10.3390/ijerph18041464.
- Brand, W., de Jongh, C., Puijkera, L., van Leeuwen, C., van Wezel, A., Schriks, M., Heringa M., van der Linden, S., Mennes, W. Establishment of trigger values and validation of bioassay panel. DEMEAU Technical Project Report (D41.2), 2014, www.demeau-fp7.eu
- Brion, F., De Gussem, V., Buchinger, S., Hollert, H., Carere, M., Porcher,J.-M., & Aït-Aïssa, S. Monitoring estrogenic activities of wasteand surface waters using a novel in vivo zebrafish embryonic (EASZY)assay: Comparison with in vitro cell-based assays and determination of effect-based trigger values. Environment International, 2019, 130, 104896.
- DB-ALM Protocol n° 197: JRC-ECVAM Automated CALUX reporter gene assay procedure
- Escher, B.I., Allinson, M., Altenburger, R., Bain, P.A., Balaguer, P., Busch, W., Crago, J., Denslow, N.D., Dopp, E., Hilscherova, K., Humpage, A.R., Anu Kumar, A., Grimaldi, M., B. Jayasinghe, S., Jarosova, B., Jia, A., Makarov, S., Maruya, K.A., Medvedev, A., Mehinto, A.C., Mendez, J.E., Poulsen, A., Prochazka, E., Richard, J., Schifferli, A., Schlenk, D., Scholz, S., Shiraishi, F., Snyder, S., Su, G., Tang, J.Y.M., van der Burg, B., van der Linden, S.C., Werner, I., Westerheide, S.D., Wong, C.K.C., Yang, M., Yeung,B.H.Y., Zhang, X., and Leusch, F.D.L. Benchmarking Organic Micropollutants in Wastewater, Recycled Water and Drinking Water with In Vitro Bioassays. Environ. Sci. Technol. 2014, 48, 3, 1940–1956.
- Escher, B. I., Aït-Aïssa, S., Behnisch, P. A., Brack, W., Brion, F., Brouwer, A., Buchinger, S., Crawford, S. E., Du Pasquier, D., Hamers, T., Hettwer, K., Hilscherová, K., Hollert, H., Kase, R., Kienle, C., Tindall, A. J., Tuerk, J., van der Oost, R., Vermeirssen, E., & Neale, P.A. Effect-based trigger values for in vitro and in vivo bioassays performed on surface water extracts supporting the environmental quality standards (EQS) of the European Water Framework Directive. Science of the Total Environment, 2018, 628–629:748–765.
- EU, European Commission Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Directive 2000/60/EC establishing a framework for Community action in the field of water policy, Directive 2006/118/EC on the protection of groundwater against pollution and deterioration and Directive 2008/105/EC on environmental quality standards in the field of water policy; Document 52022PC0540; COM/2022/544 final
- EU, Commission Implementing Decision (EU) 2018/840 of 5 June 2018 establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council and repealing Commission Implementing Decision (EU) 2015/495. 2018.; Document 32018D0840
- EU, Commission Regulation (EU) 2017/771 of 3 May 2017 amending Regulation (EC) No 152/2009 as regards the methods for the determination of the levels of dioxins and polychlorinated biphenyls; Document 32017R0771
- Gonsioroski A, Mourikes VE, Flaws JA. Endocrine Disruptors in Water and Their Effects on the Reproductive System. Int J Mol Sci. 2020 Mar 12;21(6):1929. doi: 10.3390/ijms21061929.

- ISO: International Organization for Standardization. (2018c). ISO 19040-3 – Water quality – Determination of the estrogenic potential of water and waste water – Part 3: In vitro human cell-based reporter gene assay.
- ISO: International Organization for Standardization. (2022). ISO 23196 Water quality—Calculation of biological equivalence (BEQ) concentrations.
- Kase, R., Javurkova, B., Simon, E., Swart, K., Buchinger, S., Könemann, S., Escher, B. I., Carere, M., Dulio, V., Ait-Aissa, S., Hollerr, H., Valsecchi, S., Polesello, S., Behnisch, P., di Paolo, C., Olbrich, D., Sychrova, E., Gundlach, M., Schlichting, R.,Werner, I. (2018). Screening and risk management solutions for steroidal estrogens in surface and wastewater. Trends in Analytical Chemistry, 102, 343–358.
- Kidd, K.A., Blanchfield, P.J., Mills, K.H., Palace, V.P., Evans, R.E., Lazorchak, J.M., Flick, R.W. Collapse of a fish population after exposure to a synthetic estrogen, in: Proceedings of the National Academy of Sciences, 2007, pp. 8897e8901.
- Könemann S, Kase R, Simon E, Swart K, Buchinger S, Schlüsener M, Hollert H, Escher BI, Werner I, Aït-Aïssa S, Vermeirssen E. Effect-based and chemical analytical methods to monitor estrogens under the European Water Framework Directive. TrAC Trends in Analytical Chemistry. 2018 May 1;102:225-35.
- Kunz, P.Y., Kienle, C., Carere, M., Homazava, N., & Kase, R. In vitro bioassays to screen for endocrine active pharmaceuticals in surface and waste waters. Journal of Pharmaceutical and Biomedical Analysis, 2015, 106, 107–115.
- Kunz, P.Y., Simon, E., Creusot, N., Jayasinghe, B.S., Kienle, C., Maletz, S., Schifferli, A., Schönlau, C., Selim, A-A., Denslow, N.D., Hollert, H., Werner, I., Vermeirssen, E.L.M. Effect-based tools for monitoring estrogenic mixtures: Evaluation of five in vitro bioassays, Water Research, 2017, 110, 378-388
- Loos, R., et al., 2018. Analytical methods for substances in the Watch List under the Water Framework Directive. Publications Office of the European Union.
- Loos, R., et al., 2018. in : Loos, R., et al. (Eds.), Review of the 1st Watch List under the Water Framework Directive and recommendations for the 2nd Watch List. Publication Office of the European Union.
- Lucaccioni L, Trevisani V, Marrozzini L, Bertoncelli N, Predieri B, Lugli L, Berardi A, Iughetti L. Endocrine-Disrupting Chemicals and Their Effects during Female Puberty: A Review of Current Evidence. Int J Mol Sci. 2020 Mar 18;21(6):2078. doi: 10.3390/ijms21062078.
- Neale, P.A., Escher, B.I., de Baat, M.L., Dechesne, M., Dingemans, M.M.L., Enault, J., Pronk, G.J., Smeets, P.W.M.H. and Leusch, F.D.L. Application of Effect-Based Methods to Water Quality Monitoring: Answering Frequently Asked Questions by Water Quality Managers, Regulators, and Policy Makers. Environ. Sci. Technol. 2023, 57, 15, 6023–6032
- OECD, Test No. 455: Performance-based Test Guideline for Stably Transfected Transactivation in Vitro Assays to Detect Estrogen Receptor Agonists. OECD Publishing.
- SCHER (Scientific Committee on Health and Environmental Risks), Opinion on draft environmental quality standards under the Water Framework Directive – 17β-estradiol, 30 March 2011
- Schulze T, Ahel M, Ahlheim J, Aït-Aïssa S, Brion F, Di Paolo C, Froment J, Hidasi AO, Hollender J, Hollert H, Hu M, Kloß A, Koprivica S, Krauss M, Muz M, Oswald P, Petre M, Schollée JE, Seiler TB, Shao Y, Slobodnik J, Sonavane M, Suter MJ, Tollefsen KE, Tousova Z, Walz KH, Brack W. Assessment of a novel device for onsite integrative large-volume solid phase extraction of water

samples to enable a comprehensive chemical and effect-based analysis. Sci Total Environ. 2017 Mar 1;581-582:350-358. 2017 Jan 4.

- Simon, E., Duffek, A., Stahl, C., Frey, M., Scheurer, M., Tuerk, J., Gehrmann, L., Könemann, S., Swart, K., Behnisch, P., Olbrich, D., Brion, F., Aït-Aïssa, S., Pasanen-Kase, R., Werner, I., Vermeirssen, E.L.M. Biological effect and chemical monitoring of Watch List substances in European surface waters: Steroidal estrogens and diclofenac - Effect-based methods for monitoring frameworks. Environ Int. 2022 Jan 15; 159:107033.
- Tang ZR, Xu XL, Deng SL, Lian ZX, Yu K. Oestrogenic Endocrine Disruptors in the Placenta and the Fetus. Int J Mol Sci. 2020 Feb 23;21(4):1519. doi: 10.3390/ijms21041519.
- Wagner M, Vermeirssen EL, Buchinger S, Behr M, Magdeburg A, Oehlmann J. Deriving bio-equivalents from in vitro bioassays: assessment of existing uncertainties and strategies to improve accuracy and reporting. Environ Toxicol Chem. 2013 Aug;32(8):1906-17.

A2: Vitellogenin in Male Fish as an effect-biomarker

- Ankley, G T., K M. Jensen, M. W. Hornung, M D. Kahl, J J. Korte, AND E A. Makynen. A short-term method for assessing the reproductive and developmental toxicity of endocrine-disrupting chemicals using the fathead minnow (pimephales promelas). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-01/067, 2002.
- Arukwe A, Goksøyr A. Eggshell and egg yolk proteins in fish: hepatic proteins for the next generation: oogenetic, population, and evolutionary implications of endocrine disruption. Comparative hepatology. 2003 Dec;2(1):1-21.
- Bon E, Barbe U, Rodriguez JN, Cuisset B, Pelissero C, Sumpter JP, Le Menn F. Plasma vitellogenin levels during the annual reproductive cycle of the female rainbow trout (Oncorhynchus mykiss): establishment and validation of an ELISA. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 1997 May 1;117(1):75-84.
- Brion F, Nilsen BM, Eidem JK, Goksøyr A, Porcher JM. Development and validation of an enzyme-linked immunosorbent assay to measure vitellogenin in the zebrafish (Danio rerio). Environmental Toxicology and Chemistry: An International Journal. 2002 Aug;21(8):1699-708.
- Brown RJ, Panter GH, Burden N, Salinas ER, Weltje L, Wheeler JR, Wolf Y, Lagadic L. Are changes in vitellogenin concentrations in fish reliable indicators of chemical-induced endocrine activity? Ecotoxicology and Environmental Safety. 2023 Nov 1; 266:115563.
- Burden N, Brown RJ, Holmes B, Panter GH, Salinas ER, Sewell F, Weltje L, Wheeler JR, Wolf Y, Lagadic L. An international cross-laboratory survey on fish vitellogenin analysis: Methodological challenges and opportunities for best practice. Regul Toxicol Pharmacol. 2023 Dec;145:105501. doi: 10.1016/j.yrtph.2023.105501.
- Caldwell DJ, Mastrocco F, Hutchinson TH, Länge R, Heijerick D, Janssen C, Anderson PD, Sumpter JP. Derivation of an aquatic predicted no-effect concentration for the synthetic hormone, 17 alpha-ethinyl estradiol. Environ Sci Technol. 2008 Oct 1;42(19):7046-54. doi: 10.1021/es800633q.
- Colman JR, Baldwin D, Johnson LL, Scholz NL. Effects of the synthetic estrogen, 17alpha-ethinylestradiol, on aggression and courtship behavior in male zebrafish (Danio rerio). Aquat Toxicol. 2009 Mar 9;91(4):346-54. doi: 10.1016/j.aquatox.2008.12.001.
- Cheek AO, Brouwer TH, Carroll S, Manning S, McLachlan JA, Brouwer M. Experimental evaluation of vitellogenin as a predictive biomarker for reproductive disruption. Environmental Health Perspectives. 2001 Jul;109(7):681-90.

- Dang Z. Interpretation of fish biomarker data for identification, classification, risk assessment and testing of endocrine disrupting chemicals. Environment International. 2016 Jul 1; 92:422-41.
- Fenske M, van Aerle R, Brack S, Tyler CR, Segner H. Development and validation of a homologous zebrafish (Danio rerio Hamilton-Buchanan) vitellogenin enzyme-linked immunosorbent assay (ELISA) and its application for studies on estrogenic chemicals. Comp Biochem Physiol C Toxicol Pharmacol. 2001 Jul;129(3):217-32. doi: 10.1016/s1532-0456(01)00194-6.
- Folmar LC, Gardner GR, Schreibman MP, Magliulo-Cepriano L, Mills LJ, Zaroogian G, Gutjahr-Gobell R, Haebler R, Horowitz DB, Denslow ND. Vitellogenin-induced pathology in male summer flounder (Paralichthys dentatus). Aquatic toxicology. 2001 Feb 1;51(4):431-41.
- Fujiwara Y, Fukada H, Shimizu M, Hara A. Purification of two lipovitellins and development of immunoassays for two forms of their precursors (vitellogenins) in medaka (Oryzias latipes). General and comparative endocrinology. 2005 Sep 15;143(3):267-77.
- Goodbred SL, Smith SB, Greene PS, Rauschenberger RH, Bartish TM. Reproductive and Endocrine Biomarkers in Largemouth Bass (Micropterus salmoides) and Common Carp (Cyprinus carpio) from United States Waters. US Geological Survey Data Series. 2006; 227:99.
- Hara A, Hiramatsu N, Fujita T. Vitellogenesis and choriogenesis in fishes. Fisheries science. 2016 Mar; 82:187-202.
- Hiramatsu N, Matsubara T, Fujita T, Sullivan CV, Hara A. Multiple piscine vitellogenins: biomarkers of fish exposure to estrogenic endocrine disruptors in aquatic environments. Marine Biology. 2006 Apr; 149:35-47.
- Jones PD, De Coen WM, Tremblay L, Giesy JP. Vitellogenin as a biomarker for environmental estrogens. Water Science and Technology. 2000 Oct 1;42(7-8):1-4.
- Kidd KA, Paterson MJ, Rennie MD, Podemski CL, Findlay DL, Blanchfield PJ, Liber K. Direct and indirect responses of a freshwater food web to a potent synthetic oestrogen. Philos Trans R Soc Lond B Biol Sci. 2014 Nov 19;369(1656):20130578. doi: 10.1098/rstb.2013.0578.
- Kordes C, Rieber EP, Gutzeit HO. An in vitro vitellogenin bioassay for oestrogenic substances in the medaka (Oryzias latipes). Aquatic toxicology. 2002 Aug 1;58(3-4):151-64.
- Länge R, Hutchinson TH, Croudace CP, Siegmund F, Schweinfurth H, Hampe P, Panter GH, Sumpter JP. Effects of the synthetic estrogen 17 alpha-ethinylestradiol on the life-cycle of the fathead minnow (Pimephales promelas). Environ Toxicol Chem. 2001 Jun;20(6):1216-27. doi: 10.1897/1551-5028(2001)020<1216:eotsee>2.0.co;2.
- Lahnsteiner F, Berger B, Kletzl M, Weismann T. Effect of 17beta-estradiol on gamete quality and maturation in two salmonid species. Aquat Toxicol. 2006 Aug 23;79(2):124-31. doi: 10.1016/j.aquatox.2006.05.011.
- Lee KE. Use of biological characteristics of common carp (Cyprinus carpio) to indicate exposure to hormonally active agents in selected Minnesota streams, 1999. US Department of the Interior, US Geological Survey; 2000.
- Matthiessen P, Wheeler JR, Weltje L. A review of the evidence for endocrine disrupting effects of current-use chemicals on wildlife populations. Crit Rev Toxicol. 2018 Mar;48(3):195-216. doi: 10.1080/10408444.2017.1397099.
- Parrott JL, Blunt BR. Life-cycle exposure of fathead minnows (Pimephales promelas) to an ethinylestradiol concentration below 1 ng/L reduces egg fertilization success and demasculinizes males. Environ Toxicol. 2005 Apr;20(2):131-41. doi: 10.1002/tox.20087

- Romano M, Rosanova P, Anteo C, Limatola E. Vertebrate yolk proteins: a review. Molecular Reproduction and Development: Incorporating Gamete Research. 2004 Sep;69(1):109-16.
- Schubert S, Peter A, Burki R, Schönenberger R, Suter MJ, Segner H, Burkhardt-Holm P. Sensitivity of brown trout reproduction to long-term estrogenic exposure. Aquat Toxicol. 2008 Oct 20;90(1):65-72. doi: 10.1016/j.aquatox.2008.08.002
- Scholz S, Kordes C, Hamann J, Gutzeit HO. Induction of vitellogenin in vivo and in vitro in the model teleost medaka (Oryzias latipes): comparison of gene expression and protein levels. Mar Environ Res. 2004 Apr;57(3):235-44. doi: 10.1016/S0141-1136(03)00082-5.
- Scholz S, Mayer I. Molecular biomarkers of endocrine disruption in small model fish. Mol Cell Endocrinol. 2008 Oct 10;293(1-2):57-70. doi: 10.1016/j.mce.2008.06.008.
- Solé M, Porte C, Barceló D. Analysis of the estrogenic activity of sewage treatment works and receiving waters using vitellogenin induction in fish as a biomarker. TrAC Trends in Analytical Chemistry. 2001 Sep 1;20(9):518-25.
- Sumpter JP, Jobling S. Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. Environ Health Perspect. 1995 Oct;103 Suppl 7(Suppl 7):173-8. doi: 10.1289/ehp.95103s7173.
- Tyler CR, Sumpter JP. Oocyte growth and development in teleosts. Reviews in fish biology and fisheries. 1996 Sep; 6:287-318.
- Tyler CR, Van der Eerden B, Jobling S, Panter G, Sumpter JP. Measurement of vitellogenin, a biomarker for exposure to oestrogenic chemicals, in a wide variety of cyprinid fish. Journal of Comparative Physiology B. 1996 Nov; 166:418-26.
- Valle D. Vitellogenesis in insects and other groups: a review. Memórias do Instituto Oswaldo Cruz. 1993;88:1-26.
- Wallace RA. Vitellogenesis and oocyte growth in nonmammalian vertebrates. Oogenesis. 1985 Jan 1:127-77.
- Wang YQ, Li YW, Chen QL, Liu ZH. Long-term exposure of xenoestrogens with environmental relevant concentrations disrupted spermatogenesis of zebrafish through altering sex hormone balance, stimulating germ cell proliferation, meiosis and enhancing apoptosis. Environ Pollut. 2019 Jan; 244:486-494. doi: 10.1016/j.envpol.2018.10.079.
- Watanabe KH, Jensen KM, Orlando EF, Ankley GT. What is normal? A characterization of the values and variability in reproductive endpoints of the fathead minnow, Pimephales promelas. Comp Biochem Physiol C Toxicol Pharmacol. 2007 Sep;146(3):348-56. doi: 10.1016/j.cbpc.2007.04.015.
- Zaroogian G, Gardner G, Horowitz DB, Gutjahr-Gobell R, Haebler R, Mills L. Effect of 17β-estradiol, o, p′-DDT, octylphenol and p, p′-DDE on gonadal development and liver and kidney pathology in juvenile male summer flounder (Paralichthys dentatus). Aquatic toxicology. 2001 Sep 1;54(1-2):101-12.

A3: Genotoxicity mixture threshold assessment in Occupational Biomonitoring ; the lymphocyte Cytokinesis-block micronucleus (L-CBMN) assay

- Beernaert B, Parkes EE. cGAS-STING signalling in cancer: striking a balance with chromosomal instability. Biochem Soc Trans. 2023 Apr 26;51(2):539-555. doi: 10.1042/BST20220838. PMID: 36876871; PMCID: PMC10212549.

- Bloching, M., Hofmann, A., Lautenschläger, C., Berghaus, A., & Grummt, T. Exfoliative cytology of normal buccal mucosa to predict the relative risk of cancer in the upper aerodigestive tract using the MN-assay. Oral oncology, 2000, 36(6), 550-555.
- Bolognesi, Claudia, Marco Bruzzone, Marcello Ceppi, and Francesca Marcon. "Micronuclei and upper body cancers (head, neck, breast cancers) a systematic review and meta-analysis." Mutation Research/Reviews in Mutation Research 787 (2021): 108358.
- Bonassi, Stefano, Michael Fenech, Cecilia Lando, YipingLin, Marcello Ceppi, Wushou Peter Chang,Nina Holland, et al. 2001. "Human Micronucleus Project: International Database Comparison for Results with the Cytokinesis-Block Micronucleus Assay in Human Lymphocytes: I. Effect of Laboratory Protocol, Scoring Criteria, and Host Factors on the Frequency of Micronuclei." Environmental and Molecular Mutagenesis. https://doi.org/10.1002/1098-2280(2001)37:1<31::AID-EM1004>3.0.CO;2-P.
- Bonassi S, Neri M, Lando C, Ceppi M, Lin YP, Chang WP, Holland N, Kirsch-Volders M, Zeiger E, Fenech M; HUMN collaborative group. Effect of smoking habit on the frequency of micronuclei in human lymphocytes: results from the Human MicroNucleus project. Mutat Res. 2003 Mar;543(2):155-66.
- Bonassi S, Znaor A, Ceppi M, Lando C, Chang WP, Holland N, Kirsch-Volders M, Zeiger E, Ban S, Barale R, Bigatti MP, Bolognesi C, Cebulska-Wasilewska A, Fabianova E, Fucic A, Hagmar L, Joksic G, Martelli A, Migliore L, Mirkova E, Scarfi MR, Zijno A, Norppa H, Fenech M. An increased micronucleus frequency in peripheral blood lymphocytespredicts the risk of cancer in humans. Carcinogenesis. 2007 Mar;28(3):625-31. doi: 10.1093/carcin/bgl177. Epub 2006 Sep 14. PMID: 16973674.
- Bonassi, Stefano, HannuNorppa, Marcello Ceppi, Ulf Strömberg, Roel Vermeulen, Ariana Znaor, Antonina Cebulska-Wasilewska, et al. 2008. "Chromosomal Aberration Frequency in Lymphocytes Predicts the Risk of Cancer: Results from a Pooled Cohort Study of 22 358 Subjects in 11 Countries." *Carcinogenesis*. https://doi.org/10.1093/carcin/bgn075.
- Bonassi, S., Ceppi, M., Møller, P., Azqueta, A., Milić, M., Neri, M., ... & Collins, A. (2021). DNA damage in circulating leukocytes measured with the comet assay may predict the risk of death. Scientific Reports, 11(1), 16793.
- Bonassi, S., & Fenech, M. (2021). Roadmap for translating results from the micronucleus assay into clinical practice: From observational studies to randomized controlled trials. Mutation Research/Reviews in Mutation Research, 788, 108390.
- Ceppi, M., Biasotti, B., Fenech, M., & Bonassi, S. (2010). Human population studies with the exfoliated buccal micronucleus assay: statistical and epidemiological issues. Mutation Research/Reviews in Mutation Research, 705(1), 11-19.
- CMRD 2004. DIRECTIVE 2004/37/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 29 April 2004 on the protection of workers from the risks related to exposure to carcinogens, mutagens or reprotoxic substances at work.
  Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02004L0037-20220405
- Donnellan L, Simpson B, Dhillon VS, Costabile M, Fenech M, Deo P. Methylglyoxal induces chromosomal instability and mitotic dysfunction in lymphocytes. Mutagenesis. 2021 Oct 6;36(5):339-348. doi: 10.1093/mutage/geab028. PMID: 34297102.
- Federici C, Botto N, Manfredi S, Rizza A, Del Fiandra M, Andreassi MG. Relation of increased chromosomal damage to future adverse cardiac events in patients with known coronary artery

disease. Am J Cardiol. 2008 Nov 15;102(10):1296-300. doi: 10.1016/j.amjcard.2008.07.024. Epub 2008 Sep 5. PMID: 18993144.

- Fenech M. Morley AA. Measurement of micronuclei in lymphocytes. MutatRes. 1985 Feb-Apr;147(1-2):29-36. doi: 10.1016/0165-1161(85)90015-9. PMID: 3974610.
- Fenech, M. (2000). The in vitro micronucleus technique. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 455(1-2), 81-95.
- Fenech M. Cytokinesis-block micronucleus cytome assay. Nat Protoc. 2007;2(5):1084-104. doi: 10.1038/nprot.2007.77. PMID: 17546000.
- Fenech M. Dietary reference values of individual micronutrients and nutriomes for genome damage prevention: current status and a road map to the future. Am J Clin Nutr. 2010 May;91(5):1438S-1454S. doi: 10.3945/ajcn.2010.28674D. Epub 2010 Mar 10. PMID: 20219957.
- Fenech M, Bonassi S. The effect of age, gender, diet and lifestyle on DNA damage measured using micronucleus frequency in human peripheral blood lymphocytes. Mutagenesis. 2011 Jan;26(1):43-9. doi: 10.1093/mutage/geq050. PMID: 21164181
- Fenech M. Folate (vitamin B9) and vitamin B12 and their function in the maintenance of nuclear and mitochondrial genome integrity. Mutat Res. 2012 May 1;733(1-2):21-33. doi: 10.1016/j.mrfmmm.2011.11.003. Epub 2011 Nov 7. PMID: 22093367.
- Fenech et al. 2016a. Fenech M, Knasmueller S, Bolognesi C, Bonassi S, Holland N, Migliore L, Palitti F, Natarajan AT, Kirsch-Volders M. Molecular mechanisms by which in vivo exposure to exogenous chemical genotoxic agents can lead to micronucleus formation in lymphocytes in vivo and ex vivo in humans. Mutat Res Rev Mutat Res. 2016 Oct-Dec;770(Pt A):12-25. doi: 10.1016/j.mrrev.2016.04.008. Epub 2016 Jun 7. PMID: 27894682.
- Fenech et al. 2016b. Fenech M, Nersesyan A, Knasmueller S. A systematic review of the association between occupational exposure to formaldehyde and effects on chromosomal DNA damage measured using the cytokinesis-block micronucleus assay in lymphocytes, Mutation Research/Reviews in Mutation Research,Volume 770, Part A,2016,Pages 46-57, ISSN 1383-5742,https://doi.org/10.1016/j.mrrev.2016.04.005.
- Fenech M. Cytokinesis-Block Micronucleus Cytome Assay Evolution into a More Comprehensive Method to Measure Chromosomal Instability. Genes (Basel). 2020 Oct 15;11(10):1203. doi: 10.3390/genes11101203. PMID: 33076531; PMCID: PMC7602810.
- Fenech M, Knasmueller S, Knudsen LE, Kirsch-Volders M, Deo P, Franzke B, Stopper H, Andreassi MG, Bolognesi C, Dhillon VS, Laffon B, Wagner KH, Bonassi S. "Micronuclei and Disease" special issue: Aims, scope, and synthesis of outcomes. MutatResRevMutatRes. 2021 Jul-Dec;788:108384. doi: 10.1016/j.mrrev.2021.108384. Epub 2021 Jun 5. PMID: 34893149.
- Guo X, Hintzsche H, Xu W, Ni J, Xue J, Wang X. Interplay of cGAS with micronuclei: Regulation and diseases. MutatResRevMutatRes. 2022 Jul-Dec;790:108440. doi: 10.1016/j.mrrev.2022.108440. Epub 2022 Aug 12. PMID: 35970331.
- Kirsch-Volders M, Bolognesi C, Ceppi M, Bruzzone M, Fenech M. Micronuclei, inflammation and auto-immune disease. Mutat Res Rev Mutat Res. 2020 Oct-Dec;786:108335. doi: 10.1016/j.mrrev.2020.108335. Epub 2020 Sep 15. PMID: 33339583.
- Krupina K, Goginashvili A, Cleveland DW. Scrambling the genome in cancer: causes and consequences of complex chromosome rearrangements. Nat Rev Genet. 2023 Nov 8. doi: 10.1038/s41576-023-00663-0. Epub ahead of print. PMID: 37938738.

- Mackenzie KJ, Carroll P, Martin CA, Murina O, Fluteau A, Simpson DJ, Olova N, Sutcliffe H, Rainger JK, Leitch A, Osborn RT, Wheeler AP, Nowotny M, Gilbert N, Chandra T, Reijns MAM, Jackson AP. cGAS surveillance of micronuclei links genome instability to innate immunity. Nature. 2017 Aug 24;548(7668):461-465. doi: 10.1038/nature23449. Epub 2017 Jul 24. PMID: 28738408; PMCID: PMC5870830.
- Milić, M., Ceppi, M., Bruzzone, M., Azqueta, A., Brunborg, G., Godschalk, R., ... & Bonassi, S. (2021). The hCOMET project: International database comparison of results with the comet assay in human biomonitoring. Baseline frequency of DNA damage and effect of main confounders. Mutation Research/Reviews in Mutation Research, 787, 108371.
- Møller, P., Azqueta, A., Boutet-Robinet, E., Koppen, G., Bonassi, S., Milić, M., Gajski, G., Costa, S., Teixeira, J.P., Costa Pereira, C. and Dusinska, M., 2020. Minimum Information for Reporting on the Comet Assay (MIRCA): recommendations for describing comet assay procedures and results. Nature protocols, 15(12), pp.3817-3826.
- Murgia E, Maggini V, Barale R, Rossi AM. Micronuclei, genetic polymorphisms and cardiovascular disease mortality in a nested case-control study in Italy. Mutat Res. 2007 Aug 1;621(1-2):113-8. doi: 10.1016/j.mrfmmm.2007.02.015. Epub 2007 Mar 2. PMID: 17448506.
- Murgia E, Ballardin M, Bonassi S, Rossi AM, Barale R. Validation of micronuclei frequency in peripheral blood lymphocytes as early cancer risk biomarker in a nested case-control study. Mutat Res. 2008 Mar 1;639(1-2):27-34. doi: 10.1016/j.mrfmmm.2007.10.010. Epub 2007 Nov 13. PMID: 18155071.
- Nersesyan A, Fenech M, Bolognesi C, Mišík M, Setayesh T, Wultsch G, Bonassi S, Thomas P, Knasmüller S. Use of the lymphocyte cytokinesis-block micronucleus assay in occupational biomonitoring of genome damage caused by in vivo exposure to chemical genotoxins: Past, present and future. Mutat Res Rev Mutat Res. 2016 Oct-Dec;770(Pt A):1-11. doi: 10.1016/j.mrrev.2016.05.003. Epub 2016 May 20. PMID: 27894679.
- Nersesyan et al. 2024, submitted for publication. Nersesyan A, Proietti S, Knasmueller S, Bonassi S, Fenech M. High Correlation between Micronuclei in Lymphocytes and Buccal Cells in Humans, submitted.
- NHGRI (National Human Genome Research Institute) 2016. Chromosome Abnormalities Fact Sheet. https://www.genome.gov/about-genomics/fact-sheets/Chromosome-Abnormalities-Fact-Sheet [accessed 17 March 2024].
- Rodrigues MA, Gracia García Mendoza M, Kong R, Sutton A, Pugsley HR, Li Y, Hall BE, Fogg D, Ohl L, Venkatachalam V. Automation of the Micronucleus Assay Using Imaging Flow Cytometry and Artificial Intelligence. J Vis Exp. 2023 Jan 27;(191). doi: 10.3791/64549. PMID: 36779604.
- OECD (2017). Overview of the set of OECD genetic toxicology test guidelines and updates performed in 2014–2015. Paris: Organisation for Economic Cooperation and Development (ENV/JM/MONO(2016)33/REV1; OECD Series on Testing & Assessment, No. 238, second edition; http://www.oecd.org/EHC 240: Principles for Risk Assessment of Chemicals in Food 4-112
- OECD Occupational Biomonitoring Guidance (2022), Annex Table B3, available at: https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/11/occupational-biomonitoring-guidance-document_53c29118/11bc2c7a-en.pdf
- OECD (2023). Test No. 487: In Vitro Mammalian Cell Micronucleus Test, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, https://doi.org/10.1787/9789264264861-en.

- El-Zein RA, Schabath MB, Etzel CJ, Lopez MS, Franklin JD, Spitz MR. Cytokinesis-blocked micronucleus assay as a novel biomarker for lung cancer risk. Cancer Res. 2006 Jun 15;66(12):6449-56. doi: 10.1158/0008-5472.CAN-06-0326. PMID: 16778224.
- El-Zein RA, Lopez MS, D'Amelio AM Jr, Liu M, Munden RF, Christiani D, Su L, Tejera-Alveraz P, Zhai R, Spitz MR, Etzel CJ. The cytokinesis-blocked micronucleus assay as a strong predictor of lung cancer: extension of a lung cancer risk prediction model. Cancer Epidemiol Biomarkers Prev. 2014 Nov;23(11):2462-70. doi: 10.1158/1055-9965.EPI-14-0462. Epub 2014 Aug 29. PMID: 25172871; PMCID: PMC4487661.
- Rossnerova A, Spatova M, Schunck C, Sram RJ. Automated scoring of lymphocyte micronuclei by the MetaSystemsMetafer image cytometry system and its application in studies of human mutagen sensitivity and biodosimetry of genotoxin exposure. Mutagenesis. 2011 Jan;26(1):169-75. doi: 10.1093/mutage/geq057. PMID: 21164199.
- Scarpato R, Verola C, Fabiani B, Bianchi V, Saggese G, Federico G. Nuclear damage in peripheral lymphocytes of obese and overweight Italian children as evaluated by the gamma-H2AX focus assay and micronucleus test. FASEB J. 2011 Feb;25(2):685-93. doi: 10.1096/fj.10-168427. Epub 2010 Nov 10. PMID: 21068397.
- Schunck C, Lörch T, Kowalski R, Porter M, Mahnke R, Capaccio C, Perrier J, Damer K. A standardized and automated CBMN assay for biological dosimetry: the CytoRADx™ system. RadiatProt Dosimetry. 2023 Sep 18;199(14):1516-1519. doi: 10.1093/rpd/ncad027. PMID: 37721072.
- Stopper H, Schupp N, Bahner U, Sebekova K, Klassen A, Heidland A. Genomic damage in end-stage renal failure: potential involvement of advanced glycation end products and carbonyl stress. Semin Nephrol. 2004 Sep;24(5):474-8. doi: 10.1016/j.semnephrol.2004.06.025. PMID: 15490414.
- Tavares, A.; Aimonen, K.; Ndaw, S.; Fučić, A.; Catalán, J.; Duca, R.C.; Godderis, L.; Gomes, B.C.; Janasik, B.; Ladeira, C.; et al. HBM4EU Chromates Study—Genotoxicity and Oxidative Stress Biomarkers in Workers Exposed to Hexavalent Chromium. Toxics 2022, 10, 483. https://doi.org/10.3390/toxics10080483
- Viegas S,Ladeira C, Nunes C, Malta Vacas J, Gomes M, Brito M, Mendonca P, Prista J. Genotoxic effects in occupational exposure toformaldehyde: A study in anatomy and pathologylaboratories and formaldehyde-resins production.Journal of Occupational Medicine and Toxicology 2010, 5:25. http://www.occup-med.com/content/5/1/25
- WHO (2020), Genotoxicity – Section 4.5. Available at: https://www.who.int/docs/default-source/food-safety/publications/section4-5-genotoxicity.pdf
- Wills, J. W., Verma, J. R., Rees, B. J., Harte, D. S., Haxhiraj, Q., Barnes, C. M., ... & Rees, P. (2021). Inter-laboratory automation of the in vitro micronucleus assay using imaging flow cytometry and deep learning. Archives of toxicology, 95, 3101-3115.

A4: Brain Derived Neurotrophic Factor (BDNF)assessment in serum

- Barth A., Schaffer A.W., Osterode W., Winker R., Konnaris C., Valic E., Wolf C., Rudiger H.W. Reduced cognitive abilities in lead-exposed men. Int. Arch. Occup. Environ. Health. 2002;75:394–398. doi: 10.1007/s00420-002-0329-1.
- Bellinger D.D, Very low lead exposures and children's neurodevelopment. Current Opinion in Pediatrics Volume 20, Issue 2, Pages 172 - 177April 2008
- Breno S. Diniz, Antonio L. Teixeira, Rodrigo Machado-Vieira, Leda L. Talib, Marcia Radanovic, Wagner F. Gattaz, Orestes V. Forlenza, Reduced Cerebrospinal Fluid Levels of Brain-Derived

Neurotrophic Factor Is Associated With Cognitive Impairment in Late-Life Major Depression, The Journals of Gerontology: Series B, Volume 69, Issue 6, November 2014, Pages 845–851, https://doi.org/10.1093/geronb/gbu096

- Bus B.A.A, Molendijk M.L, Penninx , B.J.W.H., Buitelaar J.K, Kenis G, Prickaerts J., Elzinga B.M, Oude Voshaar R.C. Determinants of serum brain-derived neurotrophic factor. Psychoneuroendocrinology 2011, 36 (2), p. 228-239
- Cantor AG, McDonagh MS, Blazina I, et al. Rockville (MD): Screening for Elevated Blood Lead Levels in Pregnant Women: A Systematic Review for the U.S. Preventive Services Task Force. Chapter 1. Introduction and Background. Evidence Synthesis, No. 175. Agency for Healthcare Research and Quality (US), 2019 (https://www.ncbi.nlm.nih.gov/books/NBK540426/)
- Carlino, D. et al. Low serum truncated-BDNF isoform correlates with higher cognitive impairment in schizophrenia. J Psychiatr Res 45, 273–279 (2011).
- Carlino, D., De Vanna M. & Tongiorgi, E. Is altered BDNF biosynthesis a general feature in patients with cognitive dysfunctions? Neuroscientist 19, 345–353 (2013).
- Centers for Disease Control and Prevention (CDC), 2017. Centers for Disease Control and Prevention. Biomonitoring Summary. Lead. Atlanta: CDC, 2017. https://www.cdc.gov/biomonitoring/Lead_BiomonitoringSummary.html.
- Centers for Disease Control and Prevention (CDC). CDC Response to Advisory Committee on Childhood Lead Poisoning Prevention Recommendations in “Low Level Lead Exposure Harms Children: A Renewed Call of Primary Prevention” . Atlanta (GA). 2012. [online] Available at URL: https://www.cdc.gov/nceh/lead/ACCLPP/acclpp_main.htm. 10/26/12
- Chiodo L.M, Chandice Covington, Robert J. Sokol, John H. Hannigan, James Jannise, Joel Ager, Mark Greenwald, Virginia Delaney-Black. Blood lead levels and specific attention effects in young children. Neurotoxicology and Teratology 29 (2007) 538 – 546
- Du Y, Wu HT, Qin XY, Cao C, Liu Y, Cao ZZ, Cheng Y. Postmortem Brain, Cerebrospinal Fluid, and Blood Neurotrophic Factor Levels in Alzheimer's Disease: A Systematic Review and Meta-Analysis. J Mol Neurosci. 2018 Jul;65(3):289-300. doi: 10.1007/s12031-018-1100-8. Epub 2018 Jun 28. PMID: 29956088.
- Egan MF, Kojima M, Callicott JH, Goldberg TE, Kolachana BS, Bertolino A, Zaitsev E, Gold B, Goldman D, Dean M, Lu B, Weinberger DR; The BDNF val66met Polymorphism Affects Activity-Dependent Secretion of BDNF and Human Memory and Hippocampal Function. Cell, Volume 112, Issue 2, 24 January 2003, Pages 257-269
- Fonseca-Portilla R, Krell-Roesch J, Shaibi GQ, Caselli RJ, Mandarino LJ, Zhang N, Hentz JG, Coletta DK, de Filippis EA, Dawit S, Geda YE. Brain-Derived Neurotrophic Factor and Its Associations with Metabolism and Physical Activity in a Latino Sample. Metab Syndr Relat Disord. 2019 Mar;17(2):75-80. doi: 10.1089/met.2018.0028.
- Forest M, O'Donnell KJ, Voisin G, Gaudreau H, MacIsaac JL, McEwen LM, Silveira PP, Steiner M, Kobor MS, Meaney MJ, Greenwood CMT. Agreement in DNA methylation levels from the Illumina 450K array across batches, tissues, and time. Epigenetics. 2018;13(1):19-32. doi: 10.1080/15592294.2017.1411443.
- Garcia, K. L. et al. Altered balance of proteolytic isoforms of pro-brain-derived neurotrophic factor in autism. J Neuropathol Exp Neurol 71, 289–297 (2012).
- Gao Lina, Yun Zhang, Keenan Sterling and Weihong Song. Brain-derived neurotrophic factor in Alzheimer’s disease and its pharmaceutical Potential. Translational Neurodegeneration (2022) 11:4 https://doi.org/10.1186/s40035-022-00279-0

- Gejl, A.K., Enevold, C., Bugge, A. et al. Associations between serum and plasma brain-derived neurotrophic factor and influence of storage time and centrifugation strategy. Sci Rep 9, 9655 (2019). https://doi.org/10.1038/s41598-019-45976-5
- Glass T.A., Bandeen-Roche K., McAtee M., Bolla K., Todd A.C., Schwartz B.S. Neighborhood psychosocial hazards and the association of cumulative lead dose with cognitive function in older adults. Am. J. Epidemiol. 2009;169:683–692. doi: 10.1093/aje/kwn390.
- Greiner A, Michaelsen S, Lohmann R, Weistenhofer W, Schwarz M, van Thriel C, Drexler H, Hartwig A, MAKCommission. Lead and its inorganic compounds (except lead arsenate and lead chromate) – Addendum: Evaluationof a BAT value. Assessment values in biological material – Translation of the German version from 2022. MAK Collect Occup Health Saf. 2022 Jun;7(2):Doc035. https://doi.org/10.34865/bb743992e7_2ad.
- Huang T, Gejl AK, Tarp J, Andersen LB, Peijs L, Bugge A. Cross-sectional associations of objectively measured physical activity with brain-derived neurotrophic factor in adolescents. Physiol Behav. 2017 Mar 15;171:87-91. doi: 10.1016/j.physbeh.2016.12.026.
- Huo et al., 2021, The Emerging Role of Neural Cell-Derived Exosomes in Intercellular Communication in Health and Neurodegenerative Diseases, Review Front Neurosci. 2021 Aug 31:15:738442. doi: 10.3389/fnins.2021.738442. eCollection 2021.
- Hou S., Yuan L., Jin P., Ding B., Qin N., Li L., Liu X., Wu Z., Zhao G., Deng Y. A clinical study of the effects of lead poisoning on the intelligence and neurobehavioral abilities of children. Theor. Biol. Med. Model. 2013;10:13. doi: 10.1186/1742-4682-10-13.
- Herman D.S., Geraldine M., Venkatesh T. Evaluation, diagnosis, and treatment of lead poisoning in a patient with occupational lead exposure: A case presentation. J. Occup. Med. Toxicol. 2007;2:7. doi: 10.1186/1745-6673-2-7.
- Jusko, T.A., Henderson Jr, C.R., Lanphear, B.P., Cory-Slechta, D.A., Parsons, P.J., Canfield, R.L., 2008. Blood lead concentrations< 10 μg/dL and child intelligence at 6 years of age. Environ. Health Perspect. 116 (2), 243–248
- Karri V, Schuhmacher M, Kumar V. Heavy metals (Pb, Cd, As and MeHg) as risk factors for cognitive dysfunction: A general review of metal mixture mechanism in brain. Environmental Toxicology and Pharmacology, 2016,48: 203-213.
- Klein AB, Williamson R, Santini MA, Clemmensen C, Ettrup A, Rios M, Knudsen GM, Aznar S. Blood BDNF concentrations reflect brain-tissue BDNF levels across species. Int J Neuropsychopharmacol. 2011 Apr;14(3):347-53. doi: 10.1017/S1461145710000738.
- Kundakovic M., Gudsnuk K., Herbstman J.B, Tang D., Perera F.P, and Champagne F.A. DNA methylation of BDNF as a biomarker of early-life adversity. PNAS, Neuroscience, 2014,112 (22) 6807-6813
- Legutko B, Li X, Skolnick P, 2001, Regulation of BDNF expression in primary neuron culture by LY392098, a novel AMPA receptor potentiator. Neuropharmacology, 40 (8), 1019-1027
- Li Y., Liu A., Chen K., Li L., Zhang X., Zou F., Zhang X., Meng X. Sodium butyrate alleviates lead-induced neuroinflammation and improves cognitive and memory impairment through the ACSS2/H3K9ac/ BDNF pathway. Environment International 184 (2024) 108479
- Maizlish N.A., Parra G., Feo O. Neurobehavioural evaluation of Venezuelan workers exposed to inorganic lead. Occup. Environ. Med. 1995; 52:408–414. doi: 10.1136/oem.52.6.408.

- Malavika L M, Mitra P, Goyal T, Sharma S, Purohit P, Sharma P. Association of blood lead levels with neurobehavior and BDNF expression in school going children. J Trace Elem Med Biol. 2021 Jul;66:126749. doi: 10.1016/j.jtemb.2021.126749. Epub 2021 Mar 22. PMID: 33773278.
- Mantere, P., H¨anninen, H., Hernberg, S., Luukkonen, R., 1984. A prospective follow-up study on psychological effects in workers exposed to low levels of lead. Scandinavian J. Work, Environ. Health, 10: 43–50.
- Martinowich K., Manji H., Lu B.: New insights into BDNF function in depression and anxiety. Nat. Neurosci., 10 (2007), pp. 1089-1093
- Mora E, Portella MJ, Piñol-Ripoll G, López R, Cuadras D, Forcada I, Teres M, Vieta E, Mur M. High BDNF serum levels are associated to good cognitive functioning in bipolar disorder. Eur Psychiatry. 2019 Aug;60:97-107. doi: 10.1016/j.eurpsy.2019.02.006.
- Mostafa G.A., El-Shahawi H.H., Mokhtar A. Blood lead levels in Egyptian children from high and low lead-polluted areas: Impact on cognitive function. Acta Neurol. Scand. 2009;120:30–37. doi: 10.1111/j.1600-0404.2009.01155.x.
- Muñoz Ospina B., Cadavid-Ruiz N. The effect of aerobic exercise on serum brain-derived neurotrophic factor (BDNF) and executive function in college students. Mental Health and Physical Activity. 2024, vol.26. 100578
- Mustieles V, D'Cruz SC, Couderq S, Rodríguez-Carrillo A, Fini JB, Hofer T, Steffensen IL, Dirven H, Barouki R, Olea N, Fernández MF, David A. Bisphenol A and its analogues: A comprehensive review to identify and prioritize effect biomarkers for human biomonitoring. Environ Int. 2020 Nov;144:105811. doi: 10.1016/j.envint.2020.105811. Epub 2020 Aug 28. PMID: 32866736.
- Mustieles V., Andrea Rodríguez-Carrillo, Fernando Vela-Soria, Shereen Cynthia D'Cruz, Arthur David, Fatima Smagulova, Antonio Mundo-López, Alicia Olivas-Martínez, Iris Reina-Pérez, Nicolás Olea, Carmen Freire, Juan P. Arrebola, Mariana F. Fernández. BDNF as a potential mediator between childhood BPA exposure and behavioral function in adolescent boys from the INMA-Granada cohort. Sci Total Environ. 2022 Jan 10;803:150014. doi: 10.1016/j.scitotenv.2021.150014.
- Naegelin Y, Dingsdale H, Säuberli K, Schädelin S, Kappos L, Barde YA. Measuring and Validating the Levels of Brain-Derived Neurotrophic Factor in Human Serum. eNeuro. 2018 Mar 22;5(2):ENEURO.0419-17.2018. doi: 10.1523/ENEURO.0419-17.2018. PMID: 29662942; PMCID: PMC5898630.
- Olivas-Martinez A, Suarez B, Salamanca-Fernandez E, Reina-Perez I, Rodriguez-Carrillo A, Mustieles V, Olea N, Freire C, Fernández MF. Development and validation of brain-derived neurotrophic factor measurement in human urine samples as a non-invasive effect biomarker. Front Mol Neurosci. 2023 Jan 12;15:1075613. doi: 10.3389/fnmol.2022.1075613. PMID: 36710936; PMCID: PMC9878568.
- Pedersen NH, Tarp J, Andersen LB, Gejl AK, Huang T, Peijs L, Bugge A. The association between serum brain-derived neurotrophic factor and a cluster of cardiovascular risk factors in adolescents: The CHAMPS-study DK. PLoS One. 2017 Oct 13;12(10):e0186384. doi: 10.1371/journal.pone.0186384.
- Pistollato F, Mendoza de Gyves E, Carpi D, Bopp S.K, Nunes C, Worth A, and Bal-Price A.,2020. Assessment of developmental neurotoxicity induced by chemical mixtures using an adverse outcome pathway concept. Environ Health. 2020; 19: 23. doi: 10.1186/s12940-020-00578-x.

- Polacchini A, Metelli G, Francavilla R, Baj G, Florean M, Mascaretti LG, Tongiorgi E. A method for reproducible measurements of serum BDNF: comparison of the performance of six commercial assays. Sci Rep. 2015 Dec 10;5:17989. doi: 10.1038/srep17989.
- Polyakova M, Stuke K, Schuemberg K, Mueller K, Schoenknecht P, Schroeter ML. BDNF as a biomarker for successful treatment of mood disorders: a systematic & quantitative meta-analysis. J Affect Disord. 2015 Mar 15;174:432-40. doi: 10.1016/j.jad.2014.11.044.
- Rahmani F, Saghazadeh A, Rahmani M, Teixeira AL, Rezaei N, Aghamollaii V, Ardebili HE. Plasma levels of brain-derived neurotrophic factor in patients with Parkinson disease: A systematic review and meta-analysis. Brain Res. 2019 Feb 1;1704:127-136. doi: 10.1016/j.brainres.2018.10.006.
- Ramírez Ortega D, González Esquivel DF, Blanco Ayala T, Pineda B, Gómez Manzo S, Marcial Quino J, Carrillo Mora P, Pérez de la Cruz V. Cognitive Impairment Induced by Lead Exposure during Lifespan: Mechanisms of Lead Neurotoxicity. Toxics. 2021 Jan 28;9(2):23. doi: 10.3390/toxics9020023. PMID: 33525464; PMCID: PMC7912619.
- Ren LH, Mu XY, Chen HY, Yang HL, Qi W. Prenatal lead exposure related to cord blood brain derived neurotrophic factor (BDNF) levels and impaired neonatal neurobehavioral development. Zhonghua Yu Fang Yi Xue Za Zhi. 2016 Jun;50(6):514-8. Chinese. doi: 10.3760/cma.j.issn.0253-9624.2016.06.008.
- Rodríguez-Carrillo Andrea, Verheyen Veerle J., Van Nuijs Alexander L. N., Fernández Mariana F., Remy Sylvie. Brain-derived neurotrophic factor (BDNF): an effect biomarker of neurodevelopment in human biomonitoring programs. Frontiers in Toxicology vailable ahttps://www.frontiersin.org/journals/toxicology/articles/10.3389/ftox.2023.1319788
- Sachana M, Rolaki A, Bal-Price A. Development of the Adverse Outcome Pathway (AOP): Chronic binding of antagonist to N-methyl-d-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilities of children. Toxicol Appl Pharmacol. 2018 Sep 1;354:153-175. doi: 10.1016/j.taap.2018.02.024.
- Sciarillo W.G., Alexander G., Farrell K.P. Lead exposure and child behavior. Am. J. Public Health. 1992;82:1356–1360. doi: 10.2105/AJPH.82.10.1356.
- Schwartz BS, Lee B-K, Lee G-S, Stewart WF, Lee S-S, Hwang K-Y, Ahn K-D, Kim Y-B, Bolla KI, Simon D, Parsons PJ, Todd AC (2001) Associations of blood lead, dimercaptosuccinic acid-chelatable lead, and tibia lead with neurobehavioral test scores in South Korean lead workers. Am J Epidemiol 153(5): 453–464. https://doi.org/10.1093/aje/153.5.453
- Seo J., Lee B.K., Jin S.U., Park J.W., Kim Y.T., Ryeom H.K., Lee J., Suh K.J., Kim S.H., Park S.J., et al. Lead-induced impairments in the neural processes related to working memory function. PLoS ONE. 2014;9:e105308. doi: 10.1371/journal.pone.0105308.
- Sugden K, Hannon EJ, Arseneault L, Belsky DW, Corcoran DL, Fisher HL, Houts RM, Kandaswamy R, Moffitt TE, Poulton R, Prinz JA, Rasmussen LJH, Williams BS, Wong CCY, Mill J, Caspi A. Patterns of Reliability: Assessing the Reproducibility and Integrity of DNA Methylation Measurement. Patterns (N Y). 2020 May 8;1(2):100014. doi: 10.1016/j.patter.2020.100014.
- Wei L.,He H.,1, Yang S. et al., ,Lu C. Synergistic suppression of BDNF via epigenetic mechanism deteriorating learning and memory impairment caused by Mn and Pb co-exposure. Ecotoxicology and Environmental Safety 277 (2024) 116365: https://doi.org/10.1016/j.ecoenv.2024.116365
- Weisskopf M.G., Hu H., Mulkern R.V., White R., Aro A., Oliveira S., Wright R.O. Cognitive deficits and magnetic resonance spectroscopy in adult monozygotic twins with lead poisoning. Environ. Health Perspect. 2004;112:620–625. doi: 10.1289/ehp.6687.

- White R.F., Diamond R., Proctor S., Morey C., Hu H. Residual cognitive deficits 50 years after lead poisoning during childhood. Br. J. Ind. Med. 1993;50:613–622. doi: 10.1136/oem.50.7.613.
- Winneke G., Kraemer U. Neuropsychological effects of lead in children: Interactions with social background variables. Neuropsychobiology. 1984;11:195–202. doi: 10.1159/000118077.
- Wu Z., Bai L., Tu R., Zhang L., Ba Y., Zhang H., Li X., Cheng Z., Li W., HuangH. Disruption of synaptic expression pattern and age-related DNA oxidation in a neuronal model of lead-induced toxicity. Environmental Toxicology and Pharmacology 76 (2020) 103350 doi.org/10.1016/j.etap.2020.103350
- Yalachkov Y., Anschütz V., Maiworm M., Jakob J., Schaller-Paule M.A., Schäfer J.H., Reiländer A., Friedauer L., Behrens M., Falk S., Bittner S, Foerch C. Serum and cerebrospinal fluid BDNF concentrations are associated with neurological and cognitive improvement in multiple sclerosis: A pilot study. 2023, Multiple Sclerosis and Related Disorders. 71, 104567.
- Zafra F, Lindholm D, Castren E, Hartikka J, and Thoenen H. Regulation of brain-derived neurotrophic factor and nerve growth factor mRNA in primary cultures of hippocampal neurons and astrocytes Journal of Neuroscience 1992, 12 (12) 4793-4799; DOI: https://doi.org/10.1523/JNEUROSCI.12-12-04793.1992
- Zhang B, Huo X, Xu L, Cheng Z, Cong X, Lu X, Xu X. Elevated lead levels from e-waste exposure are linked to decreased olfactory memory in children. Environ Pollut. 2017 Dec;231(Pt 1):1112-1121. doi: 10.1016/j.envpol.2017.07.015. Epub 2017 Aug 9. PMID: 28802781.
- Zhou CC, Gao ZY, He YQ, Wu MQ, Chen F, Wang J, Liu JX, Yan CH. Effects of lead, mercury, aluminium and manganese co-exposure on the serum BDNF concentration of pre-school children in Taizhou, China. Chemosphere. 2019 217:158-165).
- Zou Y., Qing L., ZengX., Shen Y., Zhong Y., Liu J., Li Q., Chen K., Lv Y., Huang D., Liang G., Zhang W., Chen L., Yang Y., Yang X., Cognitive function and plasma BDNF levels among manganese-exposed smelters. Occup Environ Med. 2014 Mar;71(3):189-94.

A5: Neurofilament light chain (NfL)in serum

- Balayssac, D., Busserolles, J., Broto, C., Dalbos, C., Prival, L., Lamoine, S., Richard, D., Quintana, M., Herbet, A., Hilairet, S., Hu, Y., Loryan, I., Glaab, W.E., Micheli, L., Ghelardini, C., Di Cesare Mannelli, L., Perrault, O., Slaoui, M., 2023. Neurofilament light chain in plasma as a sensitive diagnostic biomarker of peripheral neurotoxicity: In Vivo mouse studies with oxaliplatin and paclitaxel - NeuroDeRisk project. Biomedicine & Pharmacotherapy 167, 115535. https://doi.org/10.1016/j.biopha.2023.115535
- Barro, C., Chitnis, T., Weiner, H.L., 2020. Blood neurofilament light: a critical review of its application to neurologic disease. Annals of Clinical and Translational Neurology 7, 2508–2523. https://doi.org/10.1002/acn3.51234
- Beltran, T.A., 2024. Normative Values for Serum Neurofilament Light Chain in US Adults. Journal of Clinical Neurology 20, 46–49. https://doi.org/10.3988/jcn.2022.0340
- Benkert, P., Meier, S., Schaedelin, S., Manouchehrinia, A., Yaldizli, Ö., Maceski, A., Oechtering, J., Achtnichts, L., Conen, D., Derfuss, T., Lalive, P.H., Mueller, C., Müller, S., Naegelin, Y., Oksenberg, J.R., Pot, C., Salmen, A., Willemse, E., Kockum, I., Blennow, K., Zetterberg, H., Gobbi, C., Kappos, L., Wiendl, H., Berger, K., Sormani, M.P., Granziera, C., Piehl, F., Leppert, D., Kuhle, J., Aeschbacher, S., Barakovic, M., Buser, A., Chan, A., Disanto, G., D'Souza, M., Pasquier, R.D., Findling, O., Galbusera, R., Hrusovsky, K., Khalil, M., Lorscheider, J., Mathias, A., Orleth, A., Radue, E.-W., Rahmanzadeh, R., Sinnecker, T., Subramaniam, S., Vehoff, J., Wellmann, S.,

Wuerfel, J., Zecca, C., 2022. Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study. The Lancet Neurology 21, 246–257. https://doi.org/10.1016/S1474-4422(22)00009-6

- Beste, C., Stock, A.K., Zink, N., Ocklenburg, S., Akgün, K., Ziemssen, T., 2019. How minimal variations in neuronal cytoskeletal integrity modulate cognitive control. NeuroImage 185, 129–139. https://doi.org/10.1016/j.neuroimage.2018.10.053
- Bornhorst, J.A., Figdore, D., Campbell, M.R., Pazdernik, V.K., Mielke, M.M., Petersen, R.C., Algeciras-Schimnich, A., 2022. Plasma neurofilament light chain (NfL) reference interval determination in an Age-stratified cognitively unimpaired cohort. Clinica Chimica Acta 535, 153–156. https://doi.org/10.1016/j.cca.2022.08.017
- Flaskos, J., 2014. The neuronal cytoskeleton as a potential target in the developmental neurotoxicity of organophosphorothionate insecticides. Basic Clin Pharmacol Toxicol 115, 201–208. https://doi.org/10.1111/bcpt.12204
- Gafson, A.R., Barthélemy, N.R., Bomont, P., Carare, R.O., Durham, H.D., Julien, J.-P., Kuhle, J., Leppert, D., Nixon, R.A., Weller, R.O., Zetterberg, H., Matthews, P.M., 2020. Neurofilaments: neurobiological foundations for biomarker applications. Brain 143, 1975–1998. https://doi.org/10.1093/brain/awaa098
- Gao, Y., Su, D., Xue, Z., Ji, L., Wang, S., 2023. Association Between Serum Neurofilament Light Chain and Cognitive Performance Among Older Adults in the United States: A Cross-Sectional Study. Neurol Ther 12, 2147–2160. https://doi.org/10.1007/s40120-023-00555-9
- Khalil, M., Teunissen, C.E., Otto, M., Piehl, F., Sormani, M.P., Gattringer, T., Barro, C., Kappos, L., Comabella, M., Fazekas, F., Petzold, A., Blennow, K., Zetterberg, H., Kuhle, J., 2018. Neurofilaments as biomarkers in neurological disorders. Nature Reviews Neurology 14, 577–589. https://doi.org/10.1038/s41582-018-0058-z
- Kuhle, J., Barro, C., Andreasson, U., Derfuss, T., Lindberg, R., Sandelius, Å., Liman, V., Norgren, N., Blennow, K., Zetterberg, H., 2016. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. Clin Chem Lab Med 54, 1655–1661. https://doi.org/10.1515/cclm-2015-1195
- Rodero-Romero, A., Monreal, E., Sainz-Amo, R., García Domínguez, J.M., Villarrubia, N., Veiga-González, J.L., Fernández-Velasco, J.I., Goicochea-Briceño, H., Rodríguez-Jorge, F., Sainz de la Maza, S., Chico-García, J.L., Muriel, A., Masjuan, J., Costa-Frossard, L., Villar, L.M., 2024. Establishing Normal Serum Values of Neurofilament Light Chains and Glial Fibrillary Acidic Protein Considering the Effects of Age and Other Demographic Factors in Healthy Adults. International Journal of Molecular Sciences 25, 7808. https://doi.org/10.3390/ijms25147808
- Rosengren, L.E., Karlsson, J.E., Karlsson, J.O., Persson, L.I., Wikkelsø, C., 1996. Patients with amyotrophic lateral sclerosis and other neurodegenerative diseases have increased levels of neurofilament protein in CSF. J Neurochem 67, 2013–2018. https://doi.org/10.1046/j.1471-4159.1996.67052013.x
- Schroyen, G., Sleurs, C., Ottenbourgs, T., Leenaerts, N., Nevelsteen, I., Melis, M., Smeets, A., Deprez, S., Sunaert, S., 2023. Changes in leukoencephalopathy and serum neurofilament after (neo)adjuvant chemotherapy for breast cancer. Transl Oncol 37, 101769. https://doi.org/10.1016/j.tranon.2023.101769
- Schwartz, B.S., Lee, B.-K., Lee, G.-S., Stewart, W.F., Lee, S.-S., Hwang, K.-Y., Ahn, K.-D., Kim, Y.-B., Bolla, K.I., Simon, D., Parsons, P.J., Todd, A.C., 2001. Associations of Blood Lead, Dimercaptosuccinic Acid-chelatable Lead, and Tibia Lead with Neurobehavioral Test Scores in

South Korean Lead Workers. American Journal of Epidemiology 153, 453–464. https://doi.org/10.1093/aje/153.5.453

- Velasco, R., Argyriou, A.A., Marco, C., Mariotto, S., Stradella, A., Hernández, J., Pernas, S., Ferrari, S., Bruna, J., 2023. Serum neurofilament levels correlate with electrodiagnostic evidence of axonal loss in paclitaxel-induced peripheral neurotoxicity. J Neurol 270, 531–537. https://doi.org/10.1007/s00415-022-11377-4

A6: Neurogranin (NGRN, Ng) assessment in serum

- Abner EL, Jicha GA, Shaw LM, Trojanowski JQ, Goetzl EJ. Plasma neuronal exosomal levels of Alzheimer's disease biomarkers in normal aging. Ann Clin Transl Neurol. 2016 Apr 13;3(5):399-403. doi: 10.1002/acn3.309
- Agnello L, Lo Sasso B, Vidali M, Scazzone C, Piccoli T, Gambino CM, Bivona G, Giglio RV, Ciaccio AM, La Bella V, Ciaccio M. Neurogranin as a Reliable Biomarker for Synaptic Dysfunction in Alzheimer's Disease. Diagnostics (Basel). 2021 Dec 12;11(12):2339. doi: 10.3390/diagnostics11122339
- Alvarez XA, Winston CN, Barlow JW, Sarsoza FM, Alvarez I, Aleixandre M, Linares C, García-Fantini M, Kastberger B, Winter S, Rissman RA. Modulation of Amyloid-β and Tau in Alzheimer's Disease Plasma Neuronal-Derived Extracellular Vesicles by Cerebrolysin® and Donepezil. J Alzheimers Dis. 2022;90(2):705-717. doi: 10.3233/JAD-220575
- Bernal J. Thyroid Hormones and Brain Development. In: Pfaff DW, Arnold AP, Etgen AM, Fahrbach SE, Rubin RT (eds.). Hormones, Brain and Behavior (2nd Edition). Academic Press, 2009, pp 2005-2034, https://doi.org/10.1016/B978-008088783-8.00063-2.
- Blennow K, Zetterberg H. Biomarkers for Alzheimer's disease: current status and prospects for the future. J Intern Med. 2018 Dec;284(6):643-663. doi: 10.1111/joim.12816
- Camporesi E, Nilsson J, Brinkmalm A, Becker B, Ashton NJ, Blennow K, Zetterberg H. Fluid Biomarkers for Synaptic Dysfunction and Loss. Biomark Insights. 2020 Aug 21;15:1177271920950319. doi: 10.1177/1177271920950319
- Canturk IB, Kalkan A, Es AK, Bozan O, Unver SS, Senturk M, Ferhatlar ME, Tayfun BD. Serum Neurogranin Measurement as a Biomarker of Central Nervous System Infections: A Preliminary Study. Keio J Med. 2022 Sep 25;71(3):62-67. doi: 10.2302/kjm.2021-0019-OA
- Çevik S, Özgenç MM, Güneyk A, Evran Ş, Akkaya E, Çalış F, Katar S, Soyalp C, Hanımoğlu H, Kaynar MY. NRGN, S100B and GFAP levels are significantly increased in patients with structural lesions resulting from mild traumatic brain injuries. Clin Neurol Neurosurg. 2019 Aug;183:105380. doi: 10.1016/j.clineuro.2019.105380
- De Vos A, Jacobs D, Struyfs H, Fransen E, Andersson K, Portelius E, Andreasson U, De Surgeloose D, Hernalsteen D, Sleegers K, Robberecht C, Van Broeckhoven C, Zetterberg H, Blennow K, Engelborghs S, Vanmechelen E. C-terminal neurogranin is increased in cerebrospinal fluid but unchanged in plasma in Alzheimer's disease. Alzheimers Dement. 2015 Dec;11(12):1461-1469. doi: 10.1016/j.jalz.2015.05.012
- De Vos A, Bjerke M, Brouns R, De Roeck N, Jacobs D, Van den Abbeele L, Guldolf K, Zetterberg H, Blennow K, Engelborghs S, Vanmechelen E. Neurogranin and tau in cerebrospinal fluid and plasma of patients with acute ischemic stroke. BMC Neurol. 2017 Aug 30;17(1):170. doi: 10.1186/s12883-017-0945-8
- Dong Z, Han H, Li H, Bai Y, Wang W, Tu M, Peng Y, Zhou L, He W, Wu X, Tan T, Liu M, Wu X, Zhou W, Jin W, Zhang S, Sacktor TC, Li T, Song W, Wang YT. Long-term potentiation decay and

memory loss are mediated by AMPAR endocytosis. J Clin Invest. 2015 Jan;125(1):234-47. doi: 10.1172/JCI77888

- EU Commission 2023. “REPORT on the proposal for a directive of the European Parliament and of the Council amending Council Directive 98/24/EC and Directive 2004/37/EC as regards the limit values for lead and its inorganic compounds and diisocyanates”. European Parliament. September 8, 2023. Available at: https://www.europarl.europa.eu/doceo/document/A-9-2023-0263_EN.pdf
- Goetzl L, Darbinian N, Merabova N. Noninvasive assessment of fetal central nervous system insult: Potential application to prenatal diagnosis. Prenat Diagn. 2019 Jul;39(8):609-615. doi: 10.1002/pd.5474
- Ishitobi H, Mori K, Yoshida K, Watanabe C. Effects of perinatal exposure to low-dose cadmium on thyroid hormone-related and sex hormone receptor gene expressions in brain of offspring. Neurotoxicology. 2007 Jul;28(4):790-7. doi: 10.1016/j.neuro.2007.02.007
- Jacob S, Sumathi T. Extenuation of in utero toxic effects of MeHg in the developing neurons by Fisetin via modulating the expression of synaptic transmission and plasticity regulators in hippocampus of the rat offspring. Chem Biol Interact. 2019 May 25;305:3-10. doi: 10.1016/j.cbi.2019.03.014
- Jia L, Zhu M, Kong C, Pang Y, Zhang H, Qiu Q, Wei C, Tang Y, Wang Q, Li Y, Li T, Li F, Wang Q, Li Y, Wei Y, Jia J. Blood neuro-exosomal synaptic proteins predict Alzheimer's disease at the asymptomatic stage. Alzheimers Dement. 2021 Jan;17(1):49-60. doi: 10.1002/alz.12166
- Kalkan A, Demirel A, Atiş ŞE, Karaaslan EB, Ferhatlar ME, Senturk M. A new biomarker in the differential diagnosis of epileptic seizure: Neurogranin. Am J Emerg Med. 2022 Apr;54:147-150. doi: 10.1016/j.ajem.2022.02.010
- Krishna G, Santhoshkumar R, Sivakumar PT, Alladi S, Mahadevan A, Dahale AB, Arshad F, Subramanian S. Pathological (Dis)Similarities in Neuronal Exosome-Derived Synaptic and Organellar Marker Levels Between Alzheimer's Disease and Frontotemporal Dementia. J Alzheimers Dis. 2023;94(s1):S387-S397. doi: 10.3233/JAD-220829
- Kuşdoğan M, Vural S, Albayrak L, Çaltekin İ, Gökçen E. The diagnostic and prognostic value of serum neurogranin in acute ischemic stroke. J Stroke Cerebrovasc Dis. 2023 Feb;32(2):106889. doi: 10.1016/j.jstrokecerebrovasdis.2022.106889
- Kvartsberg H, Portelius E, Andreasson U, Brinkmalm G, Hellwig K, Lelental N, Kornhuber J, Hansson O, Minthon L, Spitzer P, Maler JM, Zetterberg H, Blennow K, Lewczuk P. Characterization of the postsynaptic protein neurogranin in paired cerebrospinal fluid and plasma samples from Alzheimer's disease patients and healthy controls. Alzheimers Res Ther. 2015 Jul 1;7(1):40. doi: 10.1186/s13195-015-0124-3
- Liu W, Lin H, He X, Chen L, Dai Y, Jia W, Xue X, Tao J, Chen L. Neurogranin as a cognitive biomarker in cerebrospinal fluid and blood exosomes for Alzheimer's disease and mild cognitive impairment. Transl Psychiatry. 2020 Apr 29;10(1):125. doi: 10.1038/s41398-020-0801-2
- Milà-Alomà M, Brinkmalm A, Ashton NJ, Kvartsberg H, Shekari M, Operto G, Salvadó G, Falcon C, Gispert JD, Vilor-Tejedor N, Arenaza-Urquijo EM, Grau-Rivera O, Sala-Vila A, Sanchez-Benavides G, González-de-Echávarri JM, Minguillon C, Fauria K, Niñerola-Baizán A, Perissinotti A, Kollmorgen G, Suridjan I, Zetterberg H, Molinuevo JL, Blennow K, Suárez-Calvet M; ALFA Study. CSF synaptic biomarkers in the preclinical stage of Alzheimer disease and their association with MRI and PET: A Cross-sectional Study. Neurology. 2021 Nov 23;97(21):e2065-e2078. doi: 10.1212/WNL.0000000000012853

- Pak JH, Huang FL, Li J, Balschun D, Reymann KG, Chiang C, Westphal H, Huang K-P. Involvement of neurogranin in the modulation of calcium/calmodulin-dependent protein kinase II, synaptic plasticity, and spatial learning: a study with knockout mice. Proc. Natl Acad. Sci. USA. 2000;97:11232–11237. https://doi.org/10.1073/pnas.210184697
- Peacock WF 4th, Van Meter TE, Mirshahi N, Ferber K, Gerwien R, Rao V, Sair HI, Diaz-Arrastia R, Korley FK. Derivation of a Three Biomarker Panel to Improve Diagnosis in Patients with Mild Traumatic Brain Injury. Front Neurol. 2017 Nov 30;8:641. doi: 10.3389/fneur.2017.00641
- Portelius E, Zetterberg H, Skillbäck T, Törnqvist U, Andreasson U, Trojanowski JQ, Weiner MW, Shaw LM, Mattsson N, Blennow K; Alzheimer's Disease Neuroimaging Initiative. Cerebrospinal fluid neurogranin: relation to cognition and neurodegeneration in Alzheimer's disease. Brain. 2015 Nov;138(Pt 11):3373-85. doi: 10.1093/brain/awv267
- Shang Y, Wang Y, Guo Y, Ren L, Zhang X, Wang S, Zhang C, Cai J. Analysis of the risk of traumatic brain injury and evaluation neurogranin and myelin basic protein as potential biomarkers of traumatic brain injury in postmortem examination. Forensic Sci Med Pathol. 2022 Sep;18(3):288-298. doi: 10.1007/s12024-022-00459-4
- Sivasangari K, Rajan KE. Standardized Bacopa monnieri Extract Ameliorates Learning and Memory Impairments through Synaptic Protein, Neurogranin, Pro-and Mature BDNF Signaling, and HPA Axis in Prenatally Stressed Rat Offspring. Antioxidants (Basel). 2020 Dec 4;9(12):1229. doi: 10.3390/antiox9121229
- Takahashi Y, Ojiro R, Yamashita R, Shimizu S, Maeda N, Zou X, Tang Q, Ozawa S, Woo GH, Yoshida T, Shibutani M. Suppression of neurogranin expression by disruption of epigenetic DNA methylation in hippocampal mature granule cells after developmental exposure to neurotoxicants in rats. Toxicol Lett. 2023 Dec 1;390:33-45. doi: 10.1016/j.toxlet.2023.10.016
- Winston CN, Goetzl EJ, Akers JC, Carter BS, Rockenstein EM, Galasko D, Masliah E, Rissman RA. Prediction of conversion from mild cognitive impairment to dementia with neuronally derived blood exosome protein profile. Alzheimers Dement (Amst). 2016 May 7;3:63-72. doi: 10.1016/j.dadm.2016.04.001
- Xiang Y, Xin J, Le W, Yang Y. Neurogranin: A Potential Biomarker of Neurological and Mental Diseases. Front Aging Neurosci. 2020 Oct 6;12:584743. doi: 10.3389/fnagi.2020.584743
- Yang J, Korley FK, Dai M, Everett AD. Serum neurogranin measurement as a biomarker of acute traumatic brain injury. Clin Biochem. 2015 Sep;48(13-14):843-8. doi: 10.1016/j.clinbiochem.2015.05.015
- Yeşilyurt Ö, Cömertpay E, Vural S, Eroğlu O, Badem ND, Çankaya İ, Bilgili YK. The diagnostic value of neurogranin in patients with carbon monoxide poisoning: Can it show early neurological damage? Am J Emerg Med. 2021 Dec;50:191-195. doi: 10.1016/j.ajem.2021.07.052

A7: Reproduction toxicity assessment in occupational biomonitoring. Male testosterone level assay

- Bhasin S, Brito JP, Cunningham GR, Hayes FJ, Hodis HN, Matsumoto AM, et al. Testosterone Therapy in Men With Hypogonadism: An Endocrine Society. Clinical Practice Guideline. The Journal of Clinical Endocrinology & Metabolism. 2018 May 1;103(5):1715–44.
- Bracci, M.; Zingaretti, L.; Martelli, M.; Lazzarini, R.; Salvio, G.; Amati, M.; Milinkovic, M.; Ulissi, A.; Medori, A.R.; Vitale, E.; et al. Alterations in Pregnenolone and Testosterone Levels in Male Shift

Workers. Int. J. Environ. Res. Public Health 2023, 20, 3195. https://doi.org/10.3390/ijerph20043195

- Clark RV, Wald JA, Swerdloff RS, et al. Large divergence in testosterone concentrations between men and women: Frame of reference for elite athletes in sex-specific competition in sports, a narrative review. Clin Endocrinol (Oxf). 2019; 90: 15–22. https://doi.org/10.1111/cen.13840
- Elizabeth G. Radke, Joseph M. Braun, John D. Meeker, Glinda S. Cooper, Phthalate exposure and male reproductive outcomes: A systematic review of the human epidemiological evidence, Environment International, Volume 121, Part 1, 2018, Pages 764-793, https://doi.org/10.1016/j.envint.2018.07.029.
- Emadi-Konjin P, Bain J, Bromberg IL. Evaluation of an algorithm for calculation of serum "Bioavailable" testosterone (BAT). Clinical Biochemistry. 2003, 36 (8): 591-596, https://doi.org/10.1016/S0009-9120(03)00076-6.
- Guzelce Ezgi Caliskan, Francesca Galbiati, Anna L. Goldman, Arijeet K. Gattu, Shehzad Basaria, Shalender Bhasin, Accurate measurement of total and free testosterone levels for the diagnosis of androgen disorders, Best Practice & Research Clinical Endocrinology & Metabolism, Volume 36, Issue 4, 2022, https://doi.org/10.1016/j.beem.2022.101683.
- Fui, Mark Ng Tang ; Dupuis, Philippe; Grossmann, Mathis. Lowered testosterone in male obesity: mechanisms, morbidity and management. Asian Journal of Andrology 16(2):p 223-231, Mar–Apr 2014. | DOI: 10.4103/1008-682X.122365
- Handelsman DJ. Androgen Physiology, Pharmacology, Use and Misuse. [Updated 2020 Oct 5]. In: Feingold KR, Anawalt B, Blackman MR, et al., editors. Endotext [Internet]. South Dartmouth (MA):MDText.com, Inc.; 2000-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK279000/
- Henrotin JB, Feigerlova E, Robert A, Dziurla M, Burgart M, Lambert-Xolin AM, Jeandel F, Weryha G. Decrease in serum testosterone levels after short-term occupational exposure to diisononyl phthalate in male workers. Occup Environ Med. 2020 Apr;77(4):214-222. https://pubmed.ncbi.nlm.nih.gov/32079716/
- Jia-Yue Zeng, Yu Miao, Chong Liu, Yan-Ling Deng, Pan-Pan Chen, Min Zhang, Fei-Peng Cui, Tian Shi, Ting-Ting Lu, Chang-Jiang Liu, Qiang Zeng, Serum multiple organochlorine pesticides in relation to testosterone concentrations among Chinese men from an infertility clinic, Chemosphere, Volume 299, 2022, https://doi.org/10.1016/j.chemosphere.2022.134469.
- Kirsten A. Baken, Nathalie Lambrechts, Sylvie Remy, Vicente Mustieles, Andrea Rodríguez-Carrillo, Christiana M. Neophytou, Nicolas Olea, Greet Schoeters, A strategy to validate a selection of human effect biomarkers using adverse outcome pathways: Proof of concept for phthalates and reproductive effects, Environmental Research, Volume 175, 2019, Pages 235-256. https://doi.org/10.1016/j.envres.2019.05.013
- Mohr BA, Guay AT, O'Donnell AB, et al. Normal, bound and nonbound testosterone levels in normally ageing men: results from the Massachusetts Male Ageing Study. Clinical endocrinology. 2005;62(1):64-73. https://doi.org/10.1111/j.1365-2265.2004.02174.x
- Pai Zheng, Zhangjian Chen, Jiaqi Shi, Yuting Xue, Yi Bai, Yulin Kang, Huiyu Xu, Guang Jia, Tiancheng Wang, Association between ambient air pollution and blood sex hormones levels in men, Environmental Research, Volume 211, 2022, 113117, https://doi.org/10.1016/j.envres.2022.113117.
- Sizar O, Leslie SW, Pico J. Androgen Replacement.. In: StatPearls [Internet]. [Updated 2023 Nov 25]. Treasure Island (FL): StatPearls Publishing; 2024 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK534853/ (Access on Sep 14, 2024)

- Sun G, Xue J, Li L, Li X, Cui Y, Qiao B, Wei D, Li H. Quantitative determination of human serum testosterone via isotope dilution ultra-performance liquid chromatography tandem mass spectrometry. Mol Med Rep. 2020 Aug;22(2):1576-1582. doi: 10.3892/mmr.2020.11235. Epub 2020 Jun 15. PMID: 32626990; PMCID: PMC7339657.
- Travison TG, Vesper HW, Orwoll E, Wu F, Kaufman JM, Wang Y, Lapauw B, Fiers T, Matsumoto AM, Bhasin S. Harmonized Reference Ranges for Circulating Testosterone Levels in Men of Four Cohort Studies in the United States and Europe. J Clin Endocrinol Metab. 2017 Apr 1;102(4):1161-1173. doi: 10.1210/jc.2016-2935. PMID: 28324103; PMCID: PMC5460736.
- Wei-Hsiang Chang, Sih-Syuan Li, Meng-Hsing Wu, Hsien-An Pan, Ching-Chang Lee, Phthalates might interfere with testicular function by reducing testosterone and insulin-like factor 3 levels, *Human Reproduction*, Volume 30, Issue 11, November 2015, Pages 2658–2670, https://doi.org/10.1093/humrep/dev225
- Wheeler MJ. The determination of bio-available testosterone. Ann Clin Biochem 1995;32:345–57.
- Xabier Arzuaga, Teneille Walker, Erin E. Yost, Elizabeth G. Radke, Andrew K. Hotchkiss, Use of the Adverse Outcome Pathway (AOP) framework to evaluate species concordance and human relevance of Dibutyl phthalate (DBP)-induced male reproductive toxicity, Reproductive Toxicology, Volume 96, 2020, Pages 445-458. https://doi.org/10.1016/j.reprotox.2019.06.009

Annex B. Effect biomarker characterization compilation

B0: Characterisation templates and answer options for ENV and HH effect biomarker assessments

Note:

Effect-biomarker characterisation leads to a knowledge of strengths and weaknesses of effect-biomarker for different aspects. The scoring is explained in the assessment templates. This characterisation allows comparability of different effect-biomarkers. This assessment concept was used in a previous guidance activity and provided evidence- based datasets. Effect biomarkers with high scores have likely better characteristics for their use. Nevertheless, like other assessments expert judgement plays a role and a simple numeric relation is not possible. For example, a low score in a critical aspect can have stronger decision impact.

List of questions for effect biomarker characterization related to environmental (ENV) assessments

Name of the effect biomarker:

Name and e-mail of the assessor(s):

Short description of the effect biomarker:

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	Non-invasive: water (5 points) hair, feathers (4 points) urine, feces samples, skin swab, mucus (3 points) Invasive:	0-5

	blood (2 points) tissue and other (1 point) None (0 points)	
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) No (0 points)	0-3
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - please report it in the comment field, add the link (3 points) No (0 points)	0-3
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	High - please provide proof of evidence via DOI or ref. in the comment field (5 points) Medium (3 points) Low (1 point) No adversity expected (0 points)	0-5
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Yes - please report it, provide the DOI in the comment field (5 points) Can be applied in a modified form (3 points) No (0 points)	0-5
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	Yes - please report it, provide the DOI (3 points) No (0 points)	0-3
7) How would you define workload, training need related to an applicability for environmental settings?	High - minimal workload and training necessary (5 points) Medium - moderate workload and training necessary (3 points) Low - high workload and expert judgement necessary (0 points) Certified commercial labs offer this bioanalysis (add 3 points)	0-8
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes - publicly available in a peer-reviewed journal. Please provide DOI in the comment field (4 points)	0-4

	Partially - as internal or non-peer-reviewed SOP (2 points) No (0 points)	
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes - OECD guideline. Please provide links in the comment field (3 points) Yes - standardized DIN EN ISO. Please provide links in the comment field (3 points) No (0 points)	0-6
10) What is the cost per sample?	Very low - <100 EURO/sample (6 points) Low - 100-250 EURO/sample (4 points) Medium - 250-400 EURO/sample (2 points) High - 400-750 EURO/sample (1 points) Very high - >750 EURO/sample (0 points)	0-6
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

List of questions for effect-biomarker characterization related to human health (HH) assessments

Name of the effect biomarker:

Name and e-mail of the assessor(s):

Short description of the effect biomarker:

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible human biological matrices?	Non-invasive: urine (5 points) saliva (4 points) buccal or nasal cells (3 points) Invasive: blood (2 points) other (1 point) None (0 points)	0-5
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) No (0 points)	0-3
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - please report it in the comment field, add the link (3 points) No (0 points)	0-3
4) Is the biomarker able to detect relevant (adverse and severe) effects in workers during a long-term exposure?	High - please provide proof of evidence via DOI or ref. in the comment field (5 points) Medium (3 points) Low (1 point) No adversity expected (0 points)	0-5
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in occupational or epidemiological studies and	Yes - please report it, provide the DOI in the comment field (5 points) Can be applied in a modified form (3 points)	0-5

resulted in meaningful results for a workplace or chemical exposure?	No (0 points)	
6) Has the biomarker been applied in environmental risk assessment or other studies with regulatory relevance (e.g. drinking water, food regulation)?	Yes - please report it, provide the DOI (3 points) No (0 points)	0-3
7) How would you define workload, training need related to an applicability for occupational settings?	High - minimal workload and training necessary (5 points) Medium - moderate workload and training necessary (3 points) Low - high workload and expert judgement necessary (0 points) Certified commercial labs offer this bioanalysis (add 3 points)	0-8
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes - publicly available in a peer reviewed journal. Please provide DOI in the comment field (4 points) Partially - as internal or non-peer-reviewed SOP (2 points) No (0 points)	0-4
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes - OECD guideline. Please provide links in the comment field (3 points) Yes - standardized DIN EN ISO. Please provide links in the comment field (3 points) No (0 points)	0-6
10) What is the cost per sample?	Very low - <100 EURO/sample (6 points) Low - 100-250 EURO/sample (4 points) Medium - 250-400 EURO/sample (2 points) High - 400-750 EURO/sample (1 points) Very high - >750 EURO/sample (0 points)	0-6
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted occupational exposure limit for a relevant reference substance?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4

12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g., does it have age dependent variations, body mass index or smoking dependency)?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4

B1: Characterisation of Estrogenic Endocrine Disruption via Estrogen Receptor alpha activity in water assessment (ENV)

Name of the effect biomarker:

Estrogen Receptor alpha (ERα) transactivation that can be assessed by *in vitro* test systems, such as the ERα-CALUX (Estrogen Receptor alpha - **C**hemical **A**ctivated **LU**ciferase gene e**X**pression) Assay

Short description of the effect-biomarker:

The ERα-CALUX is a method for the determination of the estrogenic potential of individual chemicals, chemical mixtures and/or complex environmental samples, such as water samples by means of a reporter gene assay utilizing stably transfected human cells. This reporter gene assay is based on the activation of the human estrogen receptor alpha (ISO 19040-3:2018; OECD TG455; JRC-ECVAM method 197).

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Highly specific cellular response, i.e., reporter gene expression (in human U2OS osteosarcoma cell line in the ERα-CALUX) is triggered by the receptor-binding of (xeno)estrogens. The reporter gene activation leads to an enzyme (luciferase) production and consequently to quantifiable light detection. This light signal is proportional to the amount of biological active chemicals tested or present in the sample.

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

17-beta-estradiol (E2) is the most commonly used, validated and suggested positive control (reference compound) of this assay.

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	Non-invasive: water	5
2) Is there a plausible MoA?	Yes	3
	Comment: Estrogen receptor alpha binding: Estrogen receptors (ERs) act by regulating transcriptional processes. The classical mechanism of ER action involves estrogen binding to receptors in the nucleus, after which the receptors dimerize and bind to specific response elements known as estrogen response elements (EREs) located in the promoters of target genes (Björnström and Sjöberg 2005).	

3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes	3
	Comment: The AOP for ER agonists is a fairly well-characterized AOP of the endocrine endpoints. The molecular initiating event of binding of an estrogen-mimicking chemical to the ER may be linked to adverse effects, such as reduced fertility and fecundity, altered reproduction (embryo production), with consequent effects on populations. Not all linkages are yet fully understood between the ER activation and an adverse outcome, but the ER activation appears to be a useful indicator to identify chemical exposure operating via this pathway (Gougelet et al. 2007; Kidd et al. 2007; Ankley et al. 2010; Browne et al. 2017)	
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	High	5
	Comment: Certain studies exemplify that estrogen (effect) concentrations can be quantified in various (a)biological compartments as a result of long-term exposure: • (Xeno)estrogen identification in fish bile using effect-directed analysis (Houtman et al. 2004) • Estrogenicity biomonitoring in polar extracts of sediment, suspended matter and biota (Legler et al. 2003) • Biomarker approach for investigating xenoestrogenic exposure in human blood serum (Rasmussen et al. 2003) Furthermore, the causal link between the low-concentration chronic exposure to the most potent xenoestrogen (EE2, ethinylestradiol) and the adverse effects in the fish population was clearly established in a 7-year whole-lake experiment by Kidd et al. 2007. Another study showed the (*in vivo*) ecotoxicological relevance of *in vitro*-based exposure assessments of estrogenic activity in environmental waters using the *in vitro/in vivo* comparative approach	

	(Brion et al.2019).	
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Yes Comment: Several studies are available. A few examples: • Drinking water quality assessment by effect- and chemical analyses (Villanueva et al. 2021) • Effect-based and chemical analyses of biosolids (Giudice et al. 2011) • Effect-based ozone treatment evaluation (Dopp et al. 2021)	5
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	Yes, Threshold value establishment for various water quality assessing endpoints (Escher et al. 2018; Brion et al. 2019) Regulatory chemical or water extracts screening (OECD TG455 or ISO19040:1- 3, Simon and Riegraf et al. 2022)	3
7) How would you define workload, training need related to an applicability for environmental settings?	Medium - moderate workload and training necessary	3
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes- see OECD TG455, JRC-ECVAM method 197 and ISO 19040-3	4
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes - OECD guideline OECD TG455 Yes - standardized ISO 19040-3	6
10) What is the cost per sample?	Low - 100-250 EURO/sample depending strongly on sample number to be tested. Automatization is also possible and/or the testing can be commissioned.	4
Questions for assessing sensitivity & specificity & robustness (score 0-12)		

11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	Yes. LOQs are reported to be as low as 0.002-0.2 ng/L total E2-equivalent concentration in surface- and wastewaters (Könemann et al. 2018, Simon et al. 2020, Simon and Riegraf et al. 2022)	4
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Yes – ERα-CALUX method has no relevant cross- talk to other receptors and showed high specificity, as well as sensitivity for the detection of steroidal estrogens (Kase et al. 2018; Simon et al. 2020)	4
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Yes - See outcome from EU Project New Generis project as an example: (Pedersen et al. 2010) assessing exposure to (among others) estrogenic compounds in mother-newborn pairs. The biomarker (in vitro assay) is capable of discrimininate between polluted and clean environmental sites at low pg ranges (close to LOQ concentrations): Könemann et al. 2018; Kase et al. 2018; Simon and Riegraf et al. 2022	4

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

B2: Characterisation of Estrogenic Endocrine Disruption via Vitellogenin induction in male fish (ENV)

Name of the effect biomarker:

Vitellogenin protein and/or mRNA induction in adult male fish

Short description of the effect biomarker:

Vitellogenin is an egg yolk precursor protein produced in fish and other oviparous vertebrates. It is a phosphoplipoprotein that is synthesized in the liver and transported through the blood to the ovaries where it is taken up into developing oocytes via active transport. Within oocytes vitellogenin is cleaved into multiple components including phosvitin and lipovitellin, which are incorporated into the yolk. Hepatic expression of vitellogenin is under control of estrogen response elements. Under normal conditions, reproductively mature male fish have low hepatic vtg mRNA abundance and low to non-detectable concentrations of circulating VTG protein. However, when exposed to exogenous sources of estrogen, including both natural and xenobiotic estrogenic chemicals, vitellogenin mRNA and protein can be strongly induced in males, with both mRNA abundance and protein concentrations changing by orders of magnitude. Consequently, vitellogenin in male fish is one of the most widely used and accepted biomarkers of exposure to estrogens.

While a highly sensitive and responsive biomarker of exposure, vitellogenin induction in male fish is generally not causally related to adverse effects of estrogens. Following exposure to high concentrations of estrogens, vitellogenin or is peptide products may damage the male kidney (e.g., Folmar et al. 2001; Zaroogian et al. 2001; Zha et al. 2007). However, the high levels of induction associated with kidney damage have generally only been observed in laboratory settings and are not believed to be environmentally relevant. Nonetheless, exposure to estrogen receptor agonists has repeatedly been shown to cause adverse effects on fish reproduction (e.g., Kramer et al. 1998; Matthiessen and Sumpter 1998; Kidd et al. 2007) even though vitellogenin induction in male fish may not play a causal role in the associated adverse outcome pathway(s).

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Cellular/tissue when measured as vtg mRNA transcript abundance in liver tissue

Blood parameter when VTG protein measured in plasma

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Prototypical stressors that induce vitellogenin induction in male fish include: 17β-estradiol, 17α-ethynylestradiol, diethylstilbesterol, which are all potent and highly specific estrogens.

There are many other well-known estrogenic xenobiotics including bisphenol A, 4-nonyphenol, octylphenol, but they are often weaker agonists and generally have mixed modes of action.

Question No	Answer Options	Score

Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	Vitellogenin mRNA transcripts are generally assessed in hepatic tissue sampled from animals exposed in vivo. However, vtg mRNA induction can also be assessed in liver derived cells lines. Vitellogenin protein concentrations are generally assessed in plasma. From large bodied fish, blood may be collected non-lethally. However, for small bodied fish most commonly used in toxicity testing, blood sampling is generally lethal.	2
2) Is there a plausible MoA?	Although vtg/VTG induction in males is not causally related to the adverse outcome, there is strong empirical evidence that estrogen receptor agonists can cause adverse reproductive and developmental outcomes.	3
3) Is an Adverse Outcome Pathway (AOP) reported for this effect-biomarker or is this effect-biomarker situated in an AOP network?	At present, there are no peer-reviewed and endorsed AOPs linking estrogen receptor agonism to adverse effects in vertebrates. However, there is strong empirical support in the literature that ER agonism can cause reproductive toxicity in fish.	2
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	Yes, plasma vitellogenin in fish is a fairly persistent marker of exposure to estrogens. It is cleared slowly from the plasma of male fish (who do not have a mechanism to remove it from plasma via sequestration into oocytes). Additionally, it can remain induced over long period of time. (e.g., Schmid et al. 2002; Schultz et al. 2001; Korte et al. 2009)	5
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Vitellogenin induction in male fish is most appropriately used as a biomarker of exposure. However, because the link between estrogen receptor agonism and adverse effects in vertebrates have been well established, exposure is often a very strong indicator of potential for effects. There is an extensive evidence base for the use of vitellogenin induction in male fish as a biomarker (see for example, Dang 2016).	4
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g., discharge permitting, pesticide registration, contaminated site remediation)?	Yes – the biomarker is commonly used in various regulatory guideline studies associated with endocrine disruptor screening (e.g., OECD Test No. 229; OECD Test No. 240; OECD Test No 230; OECD Test No. 234)	3

7) How would you define workload, training need related to an applicability for environmental settings?	Workload and training required to measure vtg mRNA expression is minimal. Workload and training required to measure plasma VTG protein is moderate. A number of certified commercial labs offer VTG analyses for select species. However, a species-specific antibody and vitellogenin mRNA sequences to support primer development are needed.	6
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes – publicly available in multiple peer reviewed journal articles as well as test guidelines. (e.g., Appendix B in US EPA 2002)	4
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes – see OECD Test Guidelines OECD Test No. 229; OECD Test No. 240; OECD Test No 230; OECD Test No. 234	5
10) What is the cost per sample?	Estimated cost per sample for vtg mRNA ≈15 Euro/sample Estimated cost per sample for VTG ELISA ≈20 Euro/sample – using commercial kits	6
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	Yes – the response is very sensitive and capable of responding at environmentally relevant concentrations of reference chemical.	4
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Vitellogenin induction in males is highly specific to exposure to exogenous estrogens. Other environmental variables may modulate the response somewhat, but induction is well accepted as a marker of exposure to an estrogenic stimuli. (literature too numerous to cite).	4
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	This information is readily available for comment test species (e.g., *Pimephales promelas*, *Danio rerio*, *Oryzias latipes*, *Oncorhynchus mykiss*, etc.).	4

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

References

- Dang Z. Interpretation of fish biomarker data for identification, classification, risk assessment and testing of endocrine disrupting chemicals. Environ Int. 2016 Jul-Aug;92-93:422-41. doi: 10.1016/j.envint.2016.04.003. Epub 2016 May 5. PMID: 27155823.
- Folmar LC, Gardner GR, Schreibman MP, Magliulo-Cepriano L, Mills LJ, Zaroogian G, Gutjahr-Gobell R, Haebler R, Horowitz DB, Denslow ND. Vitellogenin-induced pathology in male summer flounder (Paralichthys dentatus). Aquat Toxicol. 2001 Feb;51(4):431-41. doi: 10.1016/s0166-445x(00)00121-1.
- Kidd, K.A., Blanchfield, P.J., Mills, K.H., Palace, V.P., Evans, R.E., Lazorchak, J.M., Flick, R.W., 2007. Collapse of fish population after exposure to a synthetic estrogen. Proc. Natl. Acad. Sci. 104, 8897-8901. https://doi.org/10.1073/pnas.0609568104.
- Korte JJ, Kahl MD, Jensen KM, Pasha MS, Parks LG, LeBlanc GA, Ankley GT. Fathead minnow vitellogenin: Complementary DNA sequence and messenger RNA and protein expression after 17β-estradiol treatment. Environmental Toxicology and Chemistry: An International Journal. 2000 Apr;19(4):972-81.
- Kramer, V.J., Miles-Richardson, S., Pierens, S.L., Giesy, J.P., 1998. Reproductive impairment and
- induction of alkaline-labile phosphate, a biomarker of estrogen exposure, in fathead minnows (*Pimephales promelas*) exposed to 17β-estradiol. Aquat. Toxicol. 40, 335-360. OECD (2009), *Test No. 230: 21-day Fish Assay: A Short-Term Screening for Oestrogenic and Androgenic Activity, and Aromatase Inhibition*, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264076228-en.
- OECD (2011), *Test No. 234: Fish Sexual Development Test*, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264122369-en.
- OECD (2012), *Test No. 229: Fish Short Term Reproduction Assay*, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264185265-en.
- OECD (2015), *Test No. 240: Medaka Extended One Generation Reproduction Test (MEOGRT)*, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, https://doi.org/10.1787/9789264242258-en.
- Schmid T, Gonzalez-Valero J, Rufli H, Dietrich DR. Determination of vitellogenin kinetics in male fathead minnows (Pimephales promelas). Toxicol Lett. 2002 May 10;131(1-2):65-74. doi: 10.1016/s0378- 4274(02)00043-7. PMID: 11988359.
- Schultz IR, Orner G, Merdink JL, Skillman A. Dose-response relationships and pharmacokinetics of vitellogenin in rainbow trout after intravascular administration of 17alpha-ethynylestradiol. Aquat Toxicol. 2001 Jan;51(3):305-18. doi: 10.1016/s0166-445x(00)00118-1. PMID: 11090892.
- U.S. Environmental Protection Agency. 2002. A short-term test method for assessing the reproductive toxicity of endocrine-disrupting chemicals using the fathead minnow (*Pimephales promelas*). EPA/600/R-01/067.
- Zaroogian G, Gardner G, Horowitz DB, Gutjahr-Gobell R, Haebler R, Mills L. Effect of 17beta-estradiol, o,p'-DDT, octylphenol and p,p'-DDE on gonadal development and liver and kidney pathology in juvenile male summer flounder (Paralichthys dentatus). Aquat Toxicol. 2001 Sep;54(1-2):101-12. doi: 10.1016/s0166-445x(00)00177-6.
- Zha J, Wang Z, Wang N, Ingersoll C. Histological alternation and vitellogenin induction in adult rare minnow (Gobiocypris rarus) after exposure to ethynylestradiol and nonylphenol. Chemosphere. 2007

Jan;66(3):488-95. doi: 10.1016/j.chemosphere.2006.05.071.https://doi.org/10.1016/S0166-445X(97)00060-X.

- Matthiessen, P., Sumpter, J.P., 1998. Effects of estrogenic substances in the aquatic environment. In: Braunbeck, T., Hinton, D.E., Streit, B. (eds) Fish Ecotoxicology. EXS, vol 86. 319-335. Birkhäuser, Basel. https://doi.org/10.1007/978-3-0348-8853-0_10.

B3: Characterisation of Genotoxicity measured via micronuclei induction in humans on the example of L-CBMN (HH)

Name of the effect biomarker:

Cytokinesis-blocked micronucleus (L-CBMN) assay in human peripheral blood lymphocytes (PBL)

Short description of the effect biomarker:

The micronucleus test is used to identify substances that cause cytogenetic damage which results in the formation of micronuclei (MN). Micronuclei may originate from acentric chromosome fragments (i.e., lacking a centromere), or whole chromosomes that are unable to migrate to the poles during the anaphase stage of cell division, thus the assay detects the activity of clastogenic and aneugenic chemicals.

Used as a biomarker of early biological effect, the MN frequency in PBL is a predictive biomarker of cancer risk within a population of healthy subjects (Bonassi et al., 2007; DOI: 10.1093/carcin/bgl177). It must be noted, however, that the results should be analysed for groups of individuals as compared to a control group, for comparison of genetic damage rate between populations exposed to different environmental, occupational and lifestyle factors and population without the exposure being studied (Fenech et al., 1999; DOI: 10.1016/s1383- 5742(99)00053-8) instead of considering individual values.

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Cellular level, which is relevant for the whole organism

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Clastogens active without metabolic activation: Mitomycin B (CASRN 50-07-7); Cytosine arabinoside (147-94-4)

Aneugens: Colchicine (64-86-8); Vinblastine (143-67-9) Ionizing radiation

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	Invasive: blood (2 points)	2
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) – Genotoxicity. In addition, anti-kinetochore antibodies, FISH with pancentromeric DNA probes, or primed in situ labelling with pancentromere- specific primers, together with appropriate DNA counterstaining, can be used to identify the nature of the micronuclei (chromosome/chromosomal fragment) in order to determine whether the mechanism of micronucleus	3

	induction is due to clastogenic and/or aneugenic activity	
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - please report it in the comment field, add the link (3 points) AOP 296- Oxidative DNA damage leading to chromosomal aberrations and mutations AOP 293- Increased DNA damage leading to increased risk of breast cancer AOP 294 -Increased reactive oxygen and nitrogen species (RONS) leading to increased risk of breast cancer AOP 272 - Deposition of energy leading to lung cancer https://aopwiki.org/aops?utf8=%E2%9C%93&search=micronuclei&commit=Search&find_by_id	3
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	High - please provide proof of evidence via DOI or ref. in the comment field (5 points) Ladeira et al., 2014; DOI: 10.1080/15287394.2014.910158	5
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Yes - please report it, provide the DOI in the comment field (5 points) Ladeira et al., 2014; DOI: 10.1080/15287394.2014.910158	5
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	Yes - please report it, provide the DOI (3 points) Reviewed in Nersesyan et al., 2016; DOI: 10.1016/j.mrrev.2016.05.003	3
7) How would you define workload, training need related to an applicability for environmental settings?	Medium - moderate workload and training necessary (3 points)	3
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes - publicly available in a peer-reviewed journal. Please provide DOI in the comment field (4 points) Fenech 2007; DOI: 10.1038/nprot.2007.77 Fenech et al., 2003;	4

	DOI: 10.1016/s1383- 5718(02)00249-8	
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes - OECD guideline. Please provide links in the comment field (3 points) OECD TG 487- https://doi.org/10.1787/9789264264861-en ISO 17099:2014(en) Radiological protection — Performance criteria for laboratories using the cytokinesis block micronucleus (CBMN) assay in peripheral blood lymphocytes for biological dosimetry https://www.iso.org/standard/59141.html	3-6
10) What is the cost per sample?	Medium - 250-400 EURO/sample (2 points)	2
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	No or unknown (0 points)	0
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	No or unknown (0 points)	0
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Partially (2 points) The values have been studied in general population but show individual variations. It is always necessary to compare with a non-exposed population to conclude about the specificity of the exposure/occupational setting Fenech, 2011; doi: 10.1093/mutage/geq050.	2

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

B4: Characterisation of Oxidative stress measured via GSH/GSSG induction in humans (HH)

Name of the effect-biomarker:

Reduced/oxidized glutathione (GSH/GSSG) ratio

Short description of the effect-biomarker:

Oxidative stress level associated with carcinogenicity and genotoxicity, as well as nephrotoxicity. Glutathione is involved in several pathways related to:

1) Oxidative stress regulation that might further result in modulation of lipid peroxidation pathway (this more in relation with exposure to particles or inorganic compounds);
2) Detoxification of compounds, generally organic compounds prone to form epoxy- reactive species, since GSH will link to epoxy-compounds, which will enable an easier elimination from the body. The epoxy-compounds are known to form DNA adducts;
3) The one carbon metabolism resulting in modulation of DNA methylation.

   Glutathione is a Key Event in several AOP's:

   - AOP: 296 - Oxidative DNA damage leading to chromosomal aberrations and mutations
   - AOP: 284 - Oxidative stress in chronic kidney disease
   - AOP: 413 - Oxidation of Reduced Glutathione Leading to Mortality

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Cellular level (linked to molecular and biochemical aspects)

For the dose-response relationship, almost all reference compounds noted below are applicable. For example, lead (Pb), cadmium (Cd) and chromium (Cr), but aromatic hydrocarbons may also be of interest.

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible human biological matrices?	Noninvasive: urine (5 points) (limited number of publications). Majority of publications are on blood, but potentially urine could be used as well.	5
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) TThere is relevant literature on metals (Pb, Cd, Hg), aromatic cyclic hydrocarbons (styrene, toluene, PAHs) and other compounds that forms epoxy- metabolites (e.g., acrylamide) - a review of Rubino, 2015 about the mechanisms of glutathione-binding metals (Rubino, F.M. Toxicity of Glutathione-Binding Metals: A Review of Targets and Mechanisms. Toxics 2015, 3, 20-62. https://doi.org/10.3390/toxics3010020)	3

	- experimental evidence that GSH play a central role between detoxification, DNA methylation and DNA adducts formation (e.g. Duca et al., Scientific Reports (2018) 8:10577; DOI:10.1038/s41598-018-28911-y) thus in prolonged glutathione depletion might be related to genotoxicity and/or neurotoxicity (depending on the compounds). - a review on the toxicity of acrylamide and the evaluation of its exposure in baby foods: doi:10.1017/S0954422410000211 - a mode of action underlying development of forestomach tumors in rodents following oral exposure to ethyl acrylate and relevance to humans https://doi.org/10.1016/j.yrtph.2018.05.006	
3) Is an Adverse Outcome Pathway AOP reported for this effect-biomarker?	Yes - please report it in the comment field, add the link (3 points) Glutathione depletion is a Key Event within several AOP's: AOP 17 - Oxidative stress in developmental neurotoxicity AOP: 296 - Oxidative DNA damage leading to chromosomal aberrations and mutations AOP: 284 - Oxidative stress in chronic kidney disease AOP: 413 - Oxidation of Reduced Glutathione Leading to Mortality	3
4) Is the biomarker able to detect relevant (adverse and severe) effects in workers during a long-term exposure?	High - please provide proof of evidence via DOI or ref. in the comment field (5 points) Lead exposure: Workers: http://dx.doi.org/10.1016/j.taap.2014.10.003 http://dx.doi.org/10.1016/j.etap.2016.06.008 Styrene exposure: Workers: DOI: 10.1177/0960327111401436 https://doi.org/10.1021/tx015505x (GST polymorphism) -other articles are available, the list is not exhaustive	5
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in occupational or epidemiological studies and resulted in	Yes - please report it, provide the DOI in the comment field (5 points) Lead exposure:	5

meaningful results for a workplace or chemical exposure?	Workers: http://dx.doi.org/10.1016/j.taap.2014.10.003 http://dx.doi.org/10.1016/j.etap.2016.06.008 General Population (adults and children): doi:10.1007/s10653-017-0034-3 Styrene exposure: Workers: DOI: 10.1177/0960327111401436 https://doi.org/10.1021/tx015505x (GST polymorphism) -other articles available POPs exposure: General population (elderly population – 70 years +) http://dx.doi.org/10.1016/j.chemosphere.2014.05.013	
6) Has the biomarker been applied in environmental risk assessment or other studies with regulatory relevance (e.g., drinking water, food regulation)?	Yes - please report it, provide the DOI (3 points)	3
7) How would you define workload and applicability for occupational settings?	Medium - moderate workload and training necessary (3 points)	3
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes - publicly available in a peer-reviewed journal. Please provide DOI in the comment field (4 points) - DOI:10.1515/revac-2019-0019 (a review article with all the analytical methods) - DOI:10.1038/s41598-018-28911-y (I have a detailed SOP)	4
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect biomarker?	No (0 points)	0
10) What is the cost per sample?	Low - 100-250 EURO/sample (4 points)	4
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted occupational exposure limit for a relevant reference substance?	Yes Lead exposure: Workers: http://dx.doi.org/10.1016/j.taap.2014.10.003 Non-lead exposed workers:	4

	Blood lead levels (BLL) 4.2 ± 1.8 µg/dL GSH/GSSG 39.3 ± 6.0 Lead exposed workers: Blood lead levels (BLL) 64.8 ± 25.8 µg/dL GSH/GSSG 26.4 ± 6.8 It is to be noted that the workers (http://dx.doi.org/10.1016/j.taap.2014.10.003) have had a higher GSSG than the general population (doi:10.1007/s10653-017-0034-3). This might imply that the workers have had a metabolic stimulation due to past exposure.	
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Partially (2 points) Specificity is difficult to assess, but as stated at the beginning there are some specific mechanisms that have GSH as a common denominator. Thus, I would say this is partially specific. Could be really interesting in the context of mixture exposures to particles / inorganic compounds and aromatic hydrocarbons (which is common in industry).	2
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g., does it have age dependent variations, body mass index or smoking dependency)?	Yes - please provide reference or DOI in the comment field (4 points) General Population (adults): doi:10.1007/s10653-017-0034-3) Blood concentration; Mean ± SD; Median; Min; Max GSH (µM) 1030 ± 190; 1010; 490; 2020 GSSG (µM) 1,26 ± 1,14; 0,99; 0,25 9,59 GSH/GSSG 1200 ± 670; 1000; 100; 3220 Yes, the biomarker is sufficiently robust to different levels of exposure risks. It also has an age-dependent variation as well as a smoking dependency. However, when correcting for these co-factors, the inverse correlation between lead exposure and GSH/GSSG ratio remains significant (see.General Population article: doi:10.1007/s10653-017-0034-3)	4

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

B5: Characterisation of (Developmental) Neurotoxicity (D)NT measured via Brain Derived Neurotrophic Factor BDNF (HH)

Name of the effect biomarker:

Brain Derived Neurotrophic Factor (BDNF)

Short description of the effect-biomarker:

Alternation of BDNF is defined as a Key Event in a few currently available DNT AOPs since it plays fundamental role in brain development (neuronal survival, differentiation and morphological and functional maturation including brain plasticity). It is measured in neuronal/glial mixed culture at the protein, mRNA levels or in the cerebral-spinal fluid (CSF) as biomarker of learning and memory impairment in children. The majority of the BDNF transcripts can also be detected in the blood cells, blood serum or plasma and urine. The peripheral BDNF levels (gene expression and protein) are used as potential biomarkers for neurobehavioural performance impairment, learning and memory deficit in children and adults, psychiatric disorders such as depression (Polyakova et al., 2015), Parkinson (Rahmani et al., 2019) and Alzheimer's disease (Du et al., 2018). In addition, based on epidemiological and clinical studies BDNF DNA methylation is proposed as highly reliable effect-biomarker since it is more stable over time compared to BDNF gene expression or protein levels (Mustieles et al., 2020). In HBM4EU studies it is suggested that blood BDNF DNA methylation may be a valid surrogate marker of human brain BDNF expression levels since in both rodent and human postmortem studies data obtained from blood reflected methylation levels in the brain (especially hippocampus) (Kundakovic et al., 2015; Stenz et al., 2015). Additionally, BDNF DNA methylation at the promoter regions I and IV (the most studied) is able to predict cognitive outcomes (especially memory endpoints) among adults (Ferrer et al., 2019) and behavior in adolescents (Mustieles et al., 2022). Finally, there is some evidence of the interplay between single nucleotide polymorphisms (SNPs) of the BDNF gene and BDNF DNA methylation status on cognitive outcomes (Ferrer et al., 2019; Bakusic et al., 2021), meaning that BDNF SNPs could be used as susceptibility biomarkers to chemically-driven alterations in BDNF regulation. In this regard, the Val66Met or r6265 polymorphism is one of the most studied, being prevalent in a third- to-half of Caucasians (Shen et al., 2018).

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

BDNF levels at the cellular level (brain cells and in blood cells) and in blood serum or plasma. Recently, a validated method to measure urinary total BDNF protein levels was published (Olivas- Martínez et al., 2023).

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Lead (Pb), Bisphenol A (BPA), mercury, cadmium, arsenic

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		

1) Has the biomarker been assessed in easily accessible human biological matrices?	blood urine	2 5
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) Alternations of BDNF is defined as Key Event in several AOPs relevant to developmental neurotoxicity (DNT) and adult neurotoxicity (NT).	3
3) Is an Adverse Outcome Pathway AOP reported for this effect-biomarker?	Yes. Three AOPs relevant to DNT: AOP 12: https://aopwiki.org/aops/12 "Chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development leads to neurodegeneration with impairment in learning and memory in aging"; AOP 13: https://aopwiki.org/aops/13: "Chronic binding of antagonist to N-methyl-D-aspartate receptors (NMDARs) during brain development induces impairment of learning and memory abilities" and AOP 54: https://aopwiki.org/aops/54 "Inhibition of Na+/I-symporter (NIS) leads to learning and memory impairment" where altered BDNF level is defined as a Key Event.	3
4) Is the biomarker able to detect relevant (adverse and severe) effects in workers during a long-term exposure?	Low (2 points) Level of BDNF is relevant to both developmental and adult neurotoxicity. It is suggested as a biomarker of learning and memory impairment, psychiatric disorders and neurodegenerative disorders, including Parkinson's and Alzheimer's disease so, it could be relevant to long-term exposure to chemicals which are known to trigger neurodegeneration. Notwithstanding, there is no previous evidence of its use in occupational settings.	2
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in occupational or epidemiological studies and resulted in meaningful results for a workplace or chemical exposure?	Yes, in epidemiological studies; for instance: (1) Mustieles et al., 2022 ; (2) Zhou et al., 2019 (3) Ren et al., 2016. Further standardization of these assays is required to be used in a routine manner in occupational or epidemiological studies. BDNF DNA methylation measurements may be easier to standardize than BDNF protein levels measured with ELISA. The novel Quanterix Simoa technology can measure BDNF protein levels in serum, plasma and cerebroespinal fluid with substantially more sensitivity than ELISA kits (Ou et al., 2021).	3
6) Has the biomarker been applied in environmental risk assessment or other studies with regulatory	No (0 points)	0

relevance (e.g., drinking water, food regulation)?		
7) How would you define workload and applicability for occupational settings?	Certified commercial labs offer this bioanalysis (3 points)	3
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	BDNF levels in serum, plasma or urine are measured using commercially available ELISA kits following the detailed protocols (Polacchini et al., 2015). DNA methylation of BDNF is described in scientific publications (Kundakovic et al., 2015; Mustieles et al., 2022; Ferrer et al., 2019)	3
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect biomarker?	No (0 points)	0
10) What is the cost per sample?	Medium - 250-400 EURO/sample (2 points)) DNA methylation alone should not cost more than 100€ (although it depends on the country of measurement of course). With 400€ per sample, BDNF could be assessed in blood (methylation), serum or plasma (protein level), and urine (protein level).	2 and 6
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted occupational exposure limit for a relevant reference substance?	The LOQ for BDNF is between 7.8 and 62.5 µg/L, with Simoa technology a LOQ of 0.03 µg/L is reached and 100% of human samples can be assessed.	4
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Partially The specificity of BDNF for lead and other substances is supported by AOP13 data at a qualitative level. (https://aopwiki.org/aops/13). There is no quantitative AOP on BDNF. In the case of lead we identify BDNF as specific target.	2
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect-biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age	In adults, mean total BDNF ranges from 32.3 to 40 µg/L. In adolescents it ranges from 26.8 to 31.5 µg/L; BDNF DNA methylation: no population studies BDNF regulation can be affected by different levels of chemical exposure (Mustieles et al., 2022; Zhou et al., 2019), but also by variables such stress and childhood	4

dependent variations, body mass index or smoking dependency)?	trauma (Di Benedetto et al., 2022), physical activity (Liang et al., 2023), and alcohol intake (Cutuli et al., 2022).	

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

References:

- Bakusic J, Vrieze E, Ghosh M, Pizzagalli DA, Bekaert B, Claes S, Godderis L. Interplay of Val66Met and BDNF methylation: effect on reward learning and cognitive performance in major depression. Clin Epigenetics. 2021 Jul 29;13(1):149. doi: 10.1186/s13148-021-01136-z. Du Y, Wu HT, Qin XY, Cao C, Liu Y, Cao ZZ, Cheng Y. Postmortem Brain, Cerebrospinal Fluid, and Blood Neurotrophic Factor Levels in Alzheimer's Disease: A Systematic Review and Meta-Analysis. J Mol Neurosci. 2018 Jul;65(3):289-300. doi: 10.1007/s12031-018-1100- 8.
- Cutuli D, Sampedro-Piquero P. BDNF and its Role in the Alcohol Abuse Initiated During Early Adolescence: Evidence from Preclinical and Clinical Studies. Curr Neuropharmacol. 2022;20(11):2202-2220. doi: 10.2174/1570159X20666220624111855.
- Di Benedetto MG, Scassellati C, Cattane N, Riva MA, Cattaneo A. Neurotrophic factors, childhood trauma and psychiatric disorders: A systematic review of genetic, biochemical, cognitive and imaging studies to identify potential biomarkers. J Affect Disord. 2022 Jul 1;308:76-88. doi: 10.1016/j.jad.2022.03.071.
- Ferrer A, Labad J, Salvat-Pujol N, Barrachina M, Costas J, Urretavizcaya M, de Arriba-Arnau A, Crespo JM, Soriano-Mas C, Carracedo Á, Menchón JM, Soria V. BDNF genetic variants and methylation: effects on cognition in major depressive disorder. Transl Psychiatry. 2019 Oct 21;9(1):265. doi: 10.1038/s41398-019-0601-8.
- Kundakovic M, Gudsnuk K, Herbstman JB, Tang D, Perera FP, Champagne FA. DNA methylation of BDNF as a biomarker of early-life adversity. Proc Natl Acad Sci U S A. 2015 Jun 2;112(22):6807-13. doi: 10.1073/pnas.1408355111. Epub 2014 Nov 10. PMID:25385582; PMCID: PMC4460453.
- Liang Z, Zhang Z, Qi S, Yu J, Wei Z. Effects of a Single Bout of Endurance Exercise on Brain-Derived Neurotrophic Factor in Humans: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. Biology (Basel). 2023 Jan 13;12(1):126. doi: 10.3390/biology12010126.
- Shen T, You Y, Joseph C, Mirzaei M, Klistorner A, Graham SL, Gupta V. BDNF Polymorphism: A Review of Its Diagnostic and Clinical Relevance in Neurodegenerative Disorders. Aging Dis. 2018 Jun 1;9(3):523-536. doi: 10.14336/AD.2017.0717.
- Stenz L, Zewdie S, Laforge-Escarra T, Prados J, La Harpe R, Dayer A, Paoloni-Giacobino A, Perroud N, Aubry JM. BDNF promoter I methylation correlates between post-mortem human peripheral and brain tissues. Neurosci Res. 2015 Feb;91:1-7. doi: 10.1016/j.neures.2014.10.003.
- Mustieles V, Rodríguez-Carrillo A, Vela-Soria F, D'Cruz SC, David A, Smagulova F,Mundo-López A, Olivas-Martínez A, Reina-Pérez I, Olea N, Freire C, Arrebola JP, Fernández MF. BDNF as a potential mediator between childhood BPA exposure and behavioral function in adolescent boys from the INMA-Granada cohort. Sci Total Environ. 2022 Jan 10;803:150014. doi: 10.1016/j.scitotenv.2021.150014.

- Mustieles V, D'Cruz SC, Couderq S, Rodríguez-Carrillo A, Fini JB, Hofer T, Steffensen IL, Dirven H, Barouki R, Olea N, Fernández MF, David A. Bisphenol A and its analogues: A comprehensive review to identify and prioritize effect biomarkers for human biomonitoring. Environ Int. 2020 Nov;144:105811. doi: 10.1016/j.envint.2020.105811. Epub 2020 Aug 28. PMID: 32866736.
- Ou ZA, Byrne LM, Rodrigues FB, Tortelli R, Johnson EB, Foiani MS, Arridge M, De Vita E, Scahill RI, Heslegrave A, Zetterberg H, Wild EJ. Brain-derived neurotrophic factor in cerebrospinal fluid and plasma is not a biomarker for Huntington's disease. Sci Rep. 2021 Feb 10;11(1):3481. doi: 10.1038/s41598-021-83000-x.
- Olivas-Martinez A, Suarez B, Salamanca-Fernandez E, Reina-Perez I, Rodriguez-Carrillo A, Mustieles V, Olea N, Freire C, Fernández MF. Development and validation of brain-derived neurotrophic factor measurement in human urine samples as a non-invasive effect biomarker. Front Mol Neurosci. 2023 Jan 12;15:1075613. doi: 10.3389/fnmol.2022.1075613.
- Polacchini A, Metelli G, Francavilla R, Baj G, Florean M, Mascaretti LG, Tongiorgi E. A method for reproducible measurements of serum BDNF: comparison of the performance of six commercial assays. Sci Rep. 2015 Dec 10;5:17989. doi: 10.1038/srep17989.
- Polyakova M, Stuke K, Schuemberg K, Mueller K, Schoenknecht P, Schroeter ML. BDNF as a biomarker for successful treatment of mood disorders: a systematic & quantitative meta- analysis. J Affect Disord. 2015 Mar 15;174:432-40. doi: 10.1016/j.jad.2014.11.044.
- Rahmani F, Saghazadeh A, Rahmani M, Teixeira AL, Rezaei N, Aghamollaii V, Ardebili HE. Plasma levels of brain-derived neurotrophic factor in patients with Parkinson disease: A systematic review and meta-analysis. Brain Res. 2019 Feb 1;1704:127-136. doi: 10.1016/j.brainres.2018.10.006.
- Ren LH, Mu XY, Chen HY, Yang HL, Qi W. [Prenatal lead exposure related to cord blood brain derived neurotrophic factor (BDNF) levels and impaired neonatal neurobehavioral development]. Zhonghua Yu Fang Yi Xue Za Zhi. 2016 Jun;50(6):514-8. Chinese. doi: 10.3760/cma.j.issn.0253-9624.2016.06.008.
- Zhou CC, Gao ZY, He YQ, Wu MQ, Chen F, Wang J, Liu JX, Yan CH. Effects of lead, mercury, aluminium and manganese co-exposure on the serum BDNF concentration of pre- school children in Taizhou, China. Chemosphere. 2019 Feb;217:158-165. doi: 10.1016/j.chemosphere.2018.11.028.

B6: Characterisation of (Developmental) Neurotoxicity D)NT measured via Neurofilament light chain (NfL) (HH)

Name of the effect-biomarker:

Neuroaxonal damage/ scaffolding proteins, neurofilament light-chain (NfL) in serum (or s-NF-L)

Short description of the effect-biomarker:

Neurofilaments are the major cytoskeletal proteins in mature neurons. They are mainly expressed in axons but can also be found in cell bodies, dendrites and synapses. Their coordinated assembly is important for neurotransmission and stability as well as dynamics of synaptic compartments (Gafson et al., 2020). Mutations in the *NEFL* gene are related to peripheral neurodegeneration in Charcot-Marie-Tooth (CMT) disease. Recently, neurofilaments, especially the neurofilament light chain (NfL), have been described as biomarkers of neurological disorders (Khalil et al., 2018) and acute peripheral neurotoxicity caused by paclitaxel (Velasco et al., 2022). With the fourth generation of assays (Single molecule array (SIMOA) analysis) and a LOQ of 0.02 pg/mL it seems to be possible to investigate the integrity of the neuronal cytoskeleton of central and peripheral nerve fibers. Following axonal damage or dysfunction, NF-L leaks into the interstitial (cerebrospinal) fluid and subsequently into the bloodstream, where it can be detected and measured in serum samples. Meanwhile, z-scores for sNF-L (serum neurofilament light chain) levels are available and could be related to diseases such as multiple sclerosis (Benkert et al., 2022). The excretion pathway via the Intramural Peri-Arterial Drainage (IPAD) pathways is shared with other proteins and peptides (Albargothy et al., 2018). In addition to clinical application (mainly AD) this biomarker might be able to reflect neurobehavioral toxicity, namely the effects of sNF-L on task performance (Beste et al., 2019).

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

blood parameter

Please add suitable reference compounds which can be used as positive control in dose-response relationships:
Lead (Pb), Manganese (Mn), Aluminum (Al), Welding fumes

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	blood (2 points)	2
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) Recently, neurofilaments, especially the neurofilament light chain, has been described as biomarker of neurological disorders (Khalil et al., 2018) and acute peripheral neurotoxicity caused by paclitaxel (Velasco et al., 2022).	3

3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - Structural integrity of axons in the nervous system is measured in the KE neurite outgrowth/ integrity.	3
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	So far only effects of drugs (e.g. paclitaxel and ketamine) could be shown, current occupational studies are ongoing	2
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Can be applied in a modified form (3 points)	3
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	No (0 points)	0
7) How would you define workload, training need related to an applicability for environmental settings?	Medium - moderate workload and training necessary (3 points) Certified commercial labs offer this bioanalysis (add 3 points)	6
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes - https://doi.org/10.1038/s41582-018-0058-z	4
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	No (0 points)	0
10) What is the cost per sample?	Very low - <100 EURO/sample (6 points)	6
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	LOQ of 0.02 ng/l only for the NF-L	2
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	see Beste et al. 2019 the biomarker is Correlated stop change neurobehavioural endpoints	2
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it	see Benkert et al. 2022 for z-scores and Fitzgerald, K.C., Sotirchos, E.S., Smith, M.D., Lord, H.-N., DuVal, A., Mowry, E.M., Calabresi, P.A., 2022. Contributors to Serum NfL Levels in People without Neurologic Disease. Ann	2

have age or seasonal variations or another dependency)?	Neurol 92, 688–698. https://doi.org/10.1002/ana.26446	

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

References:

- Albargothy, N.J., Johnston, D.A., MacGregor-Sharp, M., Weller, R.O., Verma, A., Hawkes, C.A., Carare, R.O., 2018. Convective influx/glymphatic system: tracers injected into the CSF enter and leave the brain along separate periarterial basement membrane pathways. Acta Neuropathol. (Berl.) 136, 139–152. https://doi.org/10.1007/s00401-018-1862-7
- Benkert, P., Meier, S., Schaedelin, S., Manouchehrinia, A., Yaldizli, Ö., Maceski, A., Oechtering, J., Achtnichts, L., Conen, D., Derfuss, T., Lalive, P.H., Mueller, C., Müller, S., Naegelin, Y., Oksenberg, J.R., Pot, C., Salmen, A., Willemse, E., Kockum, I., Blennow, K., Zetterberg, H., Gobbi, C., Kappos, L., Wiendl, H., Berger, K., Sormani, M.P., Granziera, C., Piehl, F., Leppert, D., Kuhle, J., Aeschbacher, S., Barakovic, M., Buser, A., Chan, A., Disanto, G., D'Souza, M., Pasquier, R.D., Findling, O., Galbusera, R., Hrusovsky, K., Khalil, M., Lorscheider, J., Mathias, A., Orleth, A., Radue, E.-W., Rahmanzadeh, R., Sinnecker, T., Subramaniam, S., Vehoff, J., Wellmann, S., Wuerfel, J., Zecca, C., 2022. Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study. Lancet Neurol. 21, 246–257. https://doi.org/10.1016/S1474-4422(22)00009-6
- Beste, C., Stock, A.K., Zink, N., Ocklenburg, S., Akgün, K., Ziemssen, T., 2019. How minimal variations in neuronal cytoskeletal integrity modulate cognitive control. NeuroImage 185, 129–139. https://doi.org/10.1016/j.neuroimage.2018.10.053
- Gafson, A.R., Barthélemy, N.R., Bomont, P., Carare, R.O., Durham, H.D., Julien, J.-P., Kuhle, J., Leppert, D., Nixon, R.A., Weller, R.O., Zetterberg, H., Matthews, P.M., 2020. Neurofilaments: neurobiological foundations for biomarker applications. Brain J. Neurol. 143, 1975–1998. https://doi.org/10.1093/brain/awaa098
- Khalil, M., Teunissen, C.E., Otto, M., Piehl, F., Sormani, M.P., Gattringer, T., Barro, C., Kappos, L., Comabella, M., Fazekas, F., Petzold, A., Blennow, K., Zetterberg, H., Kuhle, J., 2018. Neurofilaments as biomarkers in neurological disorders. Nat. Rev. Neurol. 14, 577–589. https://doi.org/10.1038/s41582-018-0058-z
- Velasco, R., Argyriou, A.A., Marco, C., Mariotto, S., Stradella, A., Hernández, J., Pernas, S., Ferrari, S., Bruna, J., 2022. Serum neurofilament levels correlate with electrodiagnostic evidence of axonal loss in paclitaxel-induced peripheral neurotoxicity. J. Neurol. https://doi.org/10.1007/s00415-022-11377-4

B7: Characterisation of (Developmental) Neurotoxicity (D)NT measured via Neurogranin (Ng) (HH)

Name of the effect biomarker:

Neurogranin (Ng, NRGN)

Short description of the effect biomarker:

Neurogranin (NGRN) is a highly conserved neuron-specific postsynaptic protein specifically located in the cell bodies and dendritic processes of excitatory neurons of the cerebral cortex, hippocampus and striatum (Yang et al. 2011). In the dendritic spines, NGRN participates in the postsynaptic signalling events by regulating the availability of calmodulin (CaM) (Yang et al., 2015) and calcium mediated mechanisms (Liu et al., 2020). NGRN is involved in the cascade of events triggered by the binding of glutamate to NMDA receptors in the dendritic spines. Activation of NMDA receptors leads to a large calcium influx into postsynaptic neurons, causing an activation of a network of calcium-dependent kinases, including PKC. Activated PKC induces NGRN phosphorylation and further release of NGRN- bound CaM. Free CaM then activates its targets, such as CaM kinase II, among others (Bernal, 2009). NGRN also accelerates calcium dissociation from CaM and modulates the homeostasis of intracellular calcium in neurons (Gui et al. 2007).

NGRN has been implicated in molecular mechanisms underlying synaptogenesis, synaptic plasticity, long-term potentiation and cognitive function (Yang et al., 2015). Since synaptic function underlies cognitive performance, loss of synapses is strongly correlated with cognitive decline associated with neurodegenerative disorders (Colom-Cadena et al. 2020). An inverse correlation has been observed between NGRN levels in CSF and cognitive evaluation scores (Sanfilippo et al. 2016), such that it is considered as a CSF biomarker for synapse dysfunction and neurodegeneration (Camporesi et al. 2020).

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Molecular. So far NGRN has been assessed in CSF of patients with cognitive decline (Alzheimer's disease and other neurological and mental diseases) as it is increased in Alzheimer's disease dementia and progressive mild cognitive impairment.

NGRN can also be assessed in other biological fluids (e.g., blood). It has also been measured in foetal central nervous system-derived extracellular vesicles (NDEV) isolated from maternal plasma as early as the late first trimester.

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Not tested yet in the open literature.

Question No	Answer Options	Score

Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	blood (2 points) Several studies have shown increased NGRN levels in the CSF (Liu et al., 2020) and decreased levels in plasma neuronal-derived extracellular vesicles (NDEV) (Alvarez et al., 2022; Abner et al., 2016; Liu et al., 2020; Goetzl et al., 2016; Mouton-Liger et al. 2020; Winston et al., 2016). NGRN has also been assessed as an effect Biomarker in cord blood in an unpublished paper.	2
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) MoA: NGRN is concentrated at dendrites and spines of forebrain neurons, where it is involved in synaptic plasticity through the regulation of CaM-mediated signalling (Garrido-García et al., 2009). In neurons, NGRN binds calmodulin (CaM) and promotes a high [Ca2+]. which provides a potent signal amplification in enhancing synaptic plasticity as well as learning and memory (Huang et al., 2004). Decreased expression of neurogranin at the protein level is associated with impaired synaptic function and plasticity ultimately leading to cognitive impairments observed in Alzheimer's Disease (George et al., 2010).	3
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - please report it in the comment field, add the link (3 points) No (0 points) NGRN is mentioned in KER 1387 ("Decreased T4 in serum leads to altered hippocampal gene expression", KE:756; https://aopwiki.org/relationships/1387). Increased foetal expression of NGRN in rat hippocampus/cortex is considered as an example of empirical evidence of this KER.	1
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	Medium (3 points) NGRN as a biomarker of synaptic pathology and neurodegeneration, has the potential to detect a subtle cognitive decline in workers	3

	long-term exposed to neurotoxic chemicals (e.g., organic solvents, heavy metals, neurotoxic pesticides), which can be considered as adverse and severe effects.	
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Can be applied in a modified form (3points) NGRN has not been applied in occupational epidemiological studies published so far. It has been assessed in blood and CSF from Alzheimer's disease patients and other neurological and mental disorders (Xiang et al., 2020). Two animal studies have reported changes in NGRN expression after metal (Me-Hg and Cd) exposure. Me-Hg induced downregulation of neurogranin protein expression in the offspring rat (Jacob and Sumathi,). Perinatal cadmium exposure reduced serum thyroid hormone (TH) levels in mice pups, and as a result, decreased the NRGN gene expression, which is regulated by TH through THreceptors (Ishitobi et al., 2007)	
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	No (0 points)	0
7) How would you define workload, training need related to an applicability for environmental settings?	High - minimal workload and training necessary (5 points) Medium - moderate workload and training necessary (3 points) Low - high workload and expert judgement necessary (0 points) Certified commercial labs offer this bioanalysis (add 3 points)	3
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Partially - as internal or non-peer reviewed SOP (2 points) NRGN has been quantified in serum samples from various studies using human-specific NGRN ELISA kits (Abdel Gaber et al. 2023; Alvarez et al, 2022; De Vos et al., 2015; Kalkan et al., 2022; Kvartsberg et al., 2015; Shang et al., 2022; Yang et al., 2015; Yeşilyurt et al., 2021). A novel ultrasensitive immunoassay has been reported for measuring NGRN in CSF (Öhrfelt et al., 2020). However, it can also be measured by Luminex	2

	xMAP technology using simplex kit or multiplex panels that can be customized for users.	
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	No (0 points) NGRN is mentioned by the OECD document "OECD series on testing and assessment. Number 57. Detailed review paper on thyroid hormone disruption assays" (link). NGRN is an endpoint of thyroid hormone action. The expression of the TH-responsive gene neurogranin was significantly up-regulated by BPA in the dentate gyrus of pups following prenatal and lactational administration of BPA to rat dams (Zoeller et al., 2005)	0
10) What is the cost per sample?	Low - 100-250 EURO/sample (4 points) Invitrogen™ Neurogranin (NGRN) Human ProcartaPlex™ Simplex Kit (£214.00 / Pack of 96) (link).	4
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	Partially (2 points) LOQ (i.e., the lowest point of the calibration curve) was 10.16 pg/ml Merck Millipore (Milliplex®, sensitivity 19.9 pg/ml; standard curve range 5 -20000 pg/mL)	2
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points) As NGRN is a protein involved in the modulation of calcium/CaM signalling pathways in postsynaptic neurons, diverse chemicals and neurological and mental diseases may disrupt those pathways. Hence, NGRN is not specific for exposure to certain chemicals, but can be consider specific to assess memory consolidation and cognitive dysfunction.	0
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Partially (2 points) No or unknown (0 points) In the limited available evidence, NGRN has been shown to increase with age and is higher in females (Milà-Alomà et al., 2021). Another study found an age-dependent increase for	1

	CSF neurogranin, but not for sex (De Vos et al., 2016).	

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

Levels of NGRN in human serum samples

Median values (with interquartile range, where appropriate) or mean ± standard deviation AD: Alzheimer's disease; TBI: traumatic brain injury; MSD: Meso Scale Discovery;*Two samples collected at 3- to 11-years intervals in older cognitively intact subjects (plasma neuronal exosomal levels) ** LOD: 55 pg/ml; LOQ: 200 pg/ml (ELISA Kit); ***Data from rat serum

Author	Control	Patients	Disease
Abner et al, 2016	1st sample*: 285.1 ± 76.8 pg/mL	2nd sample*: 224.8 ± 61.0 pg/mL	Older cognitively intact subjects (two me-points)
Alvarez et al, 2022	498.61 pg/mL	288.02 pg/mL	AD
De Vos et al., 2015	1290 pg/mL (Range: 150–3780)	880 pg/mL (Range: 30–2910)	AD
Kalkan et al., 2022	979.0 pg/mL (Range: 737.1–2821.1)	1841.6 pg/mL (Range: 1101–11730)	Epilepsy
Kvartsberg et al., 2015	47451 pg/mL (MSD) (Range 21904– 90320)	36525 pg/mL (MSD) (Range 25324 to 57715)	AD
Shang et al., 2022	32.69 ± 19.61 pg/mL	139.38 ± 61.64 pg/mL	TBI
Yang et al., 2015**	20 pg/mL (IQR: 50–70 pg/mL)	180 pg/mL (IQR: 50–640 pg/mL)	TBI
Yeşilyurt et al., 2021	220 ± 100 pg/mL (Range: 40–380)	310 ± 160 pg/mL (Range: 140–980)	CO poisoning
Abdel Gaber et al. 2023***	1400 ± 50 pg/mL	670 ± 40 pg/mL	AD induced in rats

References

- Abdel Gaber SA, Hamza AH, Tantawy MA, Toraih EA, Ahmed HH. Germanium Dioxide Nanoparticles Mitigate Biochemical and Molecular Changes Characterizing Alzheimer's Disease in Rats. Pharmaceutics. 2023 Apr 30;15(5):1386. doi: 10.3390/pharmaceutics15051386
- Abner EL, Jicha GA, Shaw LM, Trojanowski JQ, Goetzl EJ. Plasma neuronal exosomal levels of Alzheimer's disease biomarkers in normal aging. Ann Clin Transl Neurol. 2016 Apr 13;3(5):399-403. doi: 10.1002/acn3.309
- Alvarez XA, Winston CN, Barlow JW, Sarsoza FM, Alvarez I, Aleixandre M, Linares C, García-Fantini M, Kastberger B, Winter S, Rissman RA. Modulation of Amyloid-β and Tau in Alzheimer's Disease Plasma Neuronal-Derived Extracellular Vesicles by Cerebrolysin® and Donepezil. J Alzheimers Dis. 2022;90(2):705-717. doi: 10.3233/JAD- 220575
- Bernal J. Thyroid Hormones and Brain Development. In: Pfaff DW, Arnold AP, Etgen AM, Fahrbach SE, Rubin RT (eds.). Hormones, Brain and Behavior (2nd Edition). Academic Press, 2009, pp 2005-2034, https://doi.org/10.1016/B978-008088783-8.00063-2
- Camporesi E, Nilsson J, Brinkmalm A, Becker B, Ashton NJ, Blennow K, Zetterberg H. Fluid Biomarkers for Synaptic Dysfunction and Loss. Biomark Insights. 2020 Aug 21;15:1177271920950319. doi: 10.1177/1177271920950319
- Colom-Cadena M, Spires-Jones T, Zetterberg H, Blennow K, Caggiano A, DeKosky ST, Fillit H, Harrison JE, Schneider LS, Scheltens P, de Haan W, Grundman M, van Dyck CH, Izzo NJ, Catalano SM; Synaptic Health Endpoints Working Group. The clinical promise of biomarkers of synapse damage or loss in Alzheimer's disease. Alzheimers Res Ther. 2020 Mar 2;12(1):21. doi: 10.1186/s13195-020-00588-4
- De Vos A, Jacobs D, Struyfs H, Fransen E, Andersson K, Portelius E, Andreasson U, De Surgeloose D, Hernalsteen D, Sleegers K, Robberecht C, Van Broeckhoven C, Zetterberg H, Blennow K, Engelborghs S, Vanmechelen E. C- terminal neurogranin is increased in cerebrospinal fluid but unchanged in plasma in Alzheimer's disease.
- Alzheimers Dement. 2015 Dec;11(12):1461-1469. doi: 10.1016/j.jalz.2015.05.012
- De Vos A, Struyfs H, Jacobs D, Fransen E, Klewansky T, De Roeck E, Robberecht C, Van Broeckhoven C, Duyckaerts C, Engelborghs S, Vanmechelen E. The Cerebrospinal Fluid Neurogranin/BACE1 Ratio is a Potential Correlate of Cognitive Decline in Alzheimer's Disease. J Alzheimers Dis. 2016 Jul 6;53(4):1523-38. doi: 10.3233/JAD-160227
- Garrido-García A, Andrés-Pans B, Durán-Trío L, Díez-Guerra FJ. Activity-dependent translocation of neurogranin to neuronal nuclei. Biochem J. 2009 Dec 10;424(3):419-29. doi: 10.1042/BJ20091071
- George AJ, Gordon L, Beissbarth T, Koukoulas I, Holsinger RM, Perreau V, Cappai R, Tan SS, Masters CL, Scott HS, Li QX. A serial analysis of gene expression profile of the Alzheimer's disease Tg2576 mouse model. Neurotox Res. 2010 May;17(4):360-79. doi: 10.1007/s12640-009-9112-3
- Goetzl EJ, Kapogiannis D, Schwartz JB, Lobach IV, Goetzl L, Abner EL, Jicha GA, Karydas AM, Boxer A, Miller BL. Decreased synaptic proteins in neuronal exosomes of frontotemporal dementia and Alzheimer's disease. FASEB J. 2016 Dec;30(12):4141-4148. doi: 10.1096/fj.201600816R
- Gui J, Song Y, Han NL, Sheu FS. Characterization of transcriptional regulation of neurogranin by nitric oxide and the role of neurogranin in SNP-induced cell death: implication of neurogranin in an

increased neuronal susceptibility to oxidative stress. Int J Biol Sci. 2007 Feb 23;3(4):212-24. doi: 10.7150/ijbs.3.212

- Huang KP, Huang FL, Jäger T, Li J, Reymann KG, Balschun D. Neurogranin/RC3 enhances long-term potentiation and learning by promoting calcium-mediated signaling. J Neurosci. 2004 Nov 24;24(47):10660-9. doi: 10.1523/JNEUROSCI.2213-04.2004
- Ishitobi H, Mori K, Yoshida K, Watanabe C. Effects of perinatal exposure to low-dose cadmium on thyroid hormone-related and sex hormone receptor gene expressions in brain of offspring. Neurotoxicology. 2007 Jul;28(4):790-7. doi: 10.1016/j.neuro.2007.02.007
- Jacob S, Sumathi T. Extenuation of in utero toxic effects of MeHg in the developing neurons by Fisetin via modulating the expression of synaptic transmission and plasticity regulators in hippocampus of the rat offspring. Chem Biol Interact. 2019 May 25;305:3-10. doi: 10.1016/j.cbi.2019.03.014
- Kalkan A, Demirel A, Atiş ŞE, Karaaslan EB, Ferhatlar ME, Senturk M. A new biomarker in the differential diagnosis of epileptic seizure: Neurogranin. Am J Emerg Med. 2022 Apr;54:147-150. doi: 10.1016/j.ajem.2022.02.010
- Kvartsberg H, Portelius E, Andreasson U, Brinkmalm G, Hellwig K, Lelental N, Kornhuber J, Hansson O, Minthon L, Spitzer P, Maler JM, Zetterberg H, Blennow K, Lewczuk P. Characterization of the postsynaptic protein neurogranin in paired cerebrospinal fluid and plasma samples from Alzheimer's disease patients and healthy controls. Alzheimers Res Ther. 2015 Jul 1;7(1):40. doi: 10.1186/s13195-015-0124-3
- Liu W, Lin H, He X, Chen L, Dai Y, Jia W, Xue X, Tao J, Chen L. Neurogranin as a cognitive biomarker in cerebrospinal fluid and blood exosomes for Alzheimer's disease and mild cognitive impairment. Transl Psychiatry. 2020 Apr 29;10(1):125. doi: 10.1038/s41398-020-0801-2
- Milà-Alomà M, Brinkmalm A, Ashton NJ, Kvartsberg H, Shekari M, Operto G, Salvadó G, Falcon C, Gispert JD, Vilor-Tejedor N, Arenaza-Urquijo EM, Grau-Rivera O, Sala-Vila A, Sanchez-Benavides G, González-de-Echávarri JM, Minguillon C, Fauria K, Niñerola-Baizán A, Perissinotti A, Kollmorgen G, Suridjan I, Zetterberg H, Molinuevo JL, Blennow K, Suárez-Calvet M; ALFA Study. CSF synaptic biomarkers in the preclinical stage of Alzheimer disease and their association with MRI and PET: A Cross-sectional Study. Neurology. 2021 Nov 23;97(21):e2065- e2078. doi: 10.1212/WNL.0000000000012853
- Mouton-Liger F, Dumurgier J, Cognat E, Hourregue C, Zetterberg H, Vanderstichele H, Vanmechelen E, Bouaziz- Amar E, Blennow K, Hugon J, Paquet C. CSF levels of the BACE1 substrate NRG1 correlate with cognition in Alzheimer's disease. Alzheimers Res Ther. 2020 Jul 20;12(1):88. doi: 10.1186/s13195-020-00655-w
- Öhrfelt A, Dumurgier J, Zetterberg H, Vrillon A, Ashton NJ, Kvartsberg H, Bouaziz-Amar E, Hugon J, Paquet C, Blennow K. Full-length and C-terminal neurogranin in Alzheimer's disease cerebrospinal fluid analyzed by novel ultrasensitive immunoassays. Alzheimers Res Ther. 2020 Dec 22;12(1):168. doi: 10.1186/s13195-020-00748-6
- Sanfilippo C, Forlenza O, Zetterberg H, Blennow K. Increased neurogranin concentrations in cerebrospinal fluid of Alzheimer's disease and in mild cognitive impairment due to AD. J Neural Transm (Vienna). 2016 Dec;123(12):1443-1447. doi: 10.1007/s00702-016-1597-3
- Shang Y, Wang Y, Guo Y, Ren L, Zhang X, Wang S, Zhang C, Cai J. Analysis of the risk of traumatic brain injury and evaluation neurogranin and myelin basic protein as potential biomarkers of traumatic brain injury in postmortem examination. Forensic Sci Med Pathol. 2022 Sep;18(3):288-298. doi: 10.1007/s12024-022-00459-4

- Winston CN, Goetzl EJ, Akers JC, Carter BS, Rockenstein EM, Galasko D, Masliah E, Rissman RA. Prediction of conversion from mild cognitive impairment to dementia with neuronally derived blood exosome protein profile. Alzheimers Dement (Amst). 2016 May 7;3:63-72. doi: 10.1016/j.dadm.2016.04.001
- Xiang Y, Xin J, Le W, Yang Y. Neurogranin: A Potential Biomarker of Neurological and Mental Diseases. Front Aging Neurosci. 2020 Oct 6;12:584743. doi: 10.3389/fnagi.2020.584743
- Yang J, Hu ZL, Jiang B, Ni L, Jin Y, Chen JG, Wang F. Effect of chloramine-T on long-term potentiation at synapses between perforant path and dentate gyrus in hippocampus of rats in vivo. Neurotoxicology. 2011 Mar;32(2):199-
- 205. doi: 10.1016/j.neuro.2011.01.006
- Yang J, Korley FK, Dai M, Everett AD. Serum neurogranin measurement as a biomarker of acute traumatic brain injury. Clin Biochem. 2015 Sep;48(13-14):843-8. doi: 10.1016/j.clinbiochem.2015.05.015
- Yeşilyurt Ö, Cömertpay E, Vural S, Eroğlu O, Badem ND, Çankaya İ, Bilgili YK. The diagnostic value of neurogranin in patients with carbon monoxide poisoning: Can it show early neurological damage? Am J Emerg Med. 2021 Dec;50:191-195. doi: 10.1016/j.ajem.2021.07.052
- Zoeller RT, Bansal R, Parris C. Bisphenol-A, an environmental contaminant that acts as a thyroid hormone receptor antagonist in vitro, increases serum thyroxine, and alters RC3/neurogranin expression in the developing rat brain. Endocrinology. 2005 Feb;146(2):607-12. doi: 10.1210/en.2004-1018

B8: Characterisation of Reproduction toxicity induced via low testosterone levels in male humans (HH)

Name of the effect-biomarker:

male testosterone level

Short description of the effect-biomarker:

Testosterone is the primary male hormone responsible for regulating sex differentiation, producing male sex characteristics, spermatogenesis and fertility. In addition, it is released into the general circulation and transported in serum, with 54% bound to sex hormone binding globulin (SHBG), 44% bound to albumin, and 2% circulating as free or unbound steroid. The free and albumin-bound fractions are considered to be biologically active. In females, its main role is as an estrogen precursor. In both genders, it also exerts anabolic effects and influences behavior [1]. Overall testosterone plays an important role in the development of the male reproductive system and is necessary for normal male fertility. Low testosterone levels are associated with adverse conditions (e.g., abnormal sexual differentiation, decreased fertility (cited in Radke et al. 2018 [2]).

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Biochemical blood parameter

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

DEHP

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	blood (2 points)	2
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) Testosterone levels are mentioned in several AOPs with plausible MoA, e.g. AOP# 18: PPARα activation in utero leading to impaired fertility in males AOP#19: Androgen receptor antagonism leading to adverse effects in the male foetus (mammals) AOP#124: HMG-CoA reductase inhibition leading to decreased fertility	3

	AOP#288: Inhibition of 17α-hydrolase/C 10,20-lyase (Cyp17A1) activity leads to birth reproductive defects (cryptorchidism) in male (mammals) (AOP#346: Aromatase inhibition leads to male-biased sex ratio via impacts on gonad differentiation AOP#376: Androgen receptor agonism leading to male-biased sex ratio) AOP#346 & 376 no Key Event related to testosterone level. Some others are listed on the AOP wiki website but are labelled «not to cite» Moreover, please have a look at the references [2-4].	
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - please report it in the comment field, add the link (3 points) See answer above and references [2-4]	3
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	Medium (3 points), see Radke et al. 2018 [2] for phthalate exposures.	3
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Can be applied in a modified form (3 points) Has been applied in several studies e.g. [2, 5-6], and resulted in meaningful results, higher exposures of different phthalates were leading to lower TT levels in exposed humans, see Radke et al. 2018 [2] (Fig.3 and Tables 17, 18, 20) for DEHP; DINP, DIBP	3
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	Yes - please report it, provide the DOI (3 points) Several studies were performed, see [7 - 10]	3
7) How would you define workload, training need related to an applicability for environmental settings?	Medium - moderate workload and training necessary (3 points) Certified commercial labs offer this bioanalysis (add 3 points)	3+3=6
Questions for assessing validation and cost (score 0-16)		

8) Does the biomarker have a well-described standard operating procedure (SOP)?	Partially - as internal or non-peer reviewed SOP (2 points), see reference [11]	2
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	No (0 points)	0
10) What is the cost per sample?	Very low - <100 EURO/sample (6 points)	6
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Usually, LOQ for TT is around 20 - 50 ng/L [11 - 12] this is by far below the normal TT levels of 40-70 years old men ranging from 158-914 ng/dL TT (1.5-9.1 µg/L) [13]	2
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Partially (2 points) The analytical specificity is high, but due to some non-occupational confounding factors (circadian rhythms, smoking, alcohol, BMI), this might become controversly discussed.	2
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Partially (2 points) TT level is age-dependent, but well investigated see [13-14]: Age 40-49:(5% / Mean/ 95%) (10.3/18.7/29) nM TT Age 50-59: (5% / Mean/ 95%) (9.0/17.3/27.7) nM TT Age 69-79: (5% / Mean/ 95%) (8.3/17.4/27.1) nM TT nM may be converted to ng/dl by dividing by 0·0347 data used from [14]	2

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

References

[1] Handelsman 2020, available at https://www.ncbi.nlm.nih.gov/books/NBK279000/#:~:text=Testosterone%20and%20its%20androgenic%20meta bolite,the%20estrogen%20receptor%20(ER)

[2] Radke et al. 2018, available at https://www.sciencedirect.com/science/article/pii/S0160412018303404?via%3Dihub

[3] – Baken et al 2019, available at https://doi.org/10.1016/j.envres.2019.05.013

[4] – Arzuaga el 2020, available at https://doi.org/10.1016/j.reprotox.2019.06.009

[5] - Henrotin J-B, et al. 2020, available at: https://pubmed.ncbi.nlm.nih.gov/32079716/

[6] – Bracci et al 2023, available at https://doi.org/10.3390/ijerph20043195

[7] - Chang et al. 2015, available at https://academic.oup.com/humrep/article/30/11/2658/2385019

[8] - Zeng et al. 2022, available at https://doi.org/10.1016/j.chemosphere.2022.134469

[9] - Kimbald et al 2022, available at https://doi.org/10.1111/andr.13198

[10] - Zheng et al. 2022, available at https://doi.org/10.1016/j.envres.2022.113117

[11] - Sun et al., Mol Med Rep. 2020doi: 10.3892/mmr.2020.11235; available at: https://www.spandidos- publications.com/10.3892/mmr.2020.11235

[12] - Matsumuto et al. 2004, available at https://doi.org/10.1210/jc.2003-032175

[13] – Mohr at al. 2004, available at https://doi.org/10.1111/j.1365-2265.2004.02174.x

[14] - Clark et al 2018, available at https://onlinelibrary.wiley.com/doi/epdf/10.1111/cen.13840

B9: Characterisation of Genotoxicity measured via Comet assay with frozen samples (HH)

Name of the effect biomarker:

Comet assay with frozen samples

Short description of the effect-biomarker:

The Comet assay is a sensitive, inexpensive, rapid, and easy-to-use method that is widely used in genotoxicity testing and biomonitoring studies. It can be applied using a few cells and on almost every cell type to assess a range of different DNA lesions, DNA repair activity and even DNA methylation levels, at the individual cell level. The principle is simple, during electrophoresis relaxed DNA strands move away from the nucleoid core and towards the anode, appearing like a “comet”, hence the name of the assay. The distribution of DNA within the comet relates to the extent of DNA damage present in the cell. The Comet assay presents numerous advantages including: (1) sensitivity for detecting low levels of damage, (2) use of any monodispersed cell population, proliferating as well as nonproliferating, (3) single cell data collection, allowing more robust statistical analyses, (4) requirement for a small number of cells per sample, (5) low cost, rapid, and easy application, and (6) flexibility to use fresh or frozen samples. In fact, any eukaryote cell that can be obtained as a single cell suspension can potentially be used. Well-characterized cell lines or primary cells (e.g., peripheral blood mononuclear cells, buccal cells, cells of tissue biopsies, placenta) retrieved from different biological matrices such as blood, saliva, tears or sperm cells have been used in biomonitoring studies. An extra step after lysis, digestion with lesion-specific enzymes, namely, DNA glycosylases has improved Comet assay’s sensitivity and selectivity, being possible to measure oxidative DNA damage and DNA repair.

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Molecular and Biochemical

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

H_2O_2- single breaks $KBrO_3$- oxidative damage

Question No	Answer Options	Score	Evaluation
Questions for assessing relevance and invasiveness (score 0-16)			
1) Has the biomarker been assessed in easily accessible matrices?	Non-invasive: water (5 points) hair, feathers (4 points) urine, feces samples, skin swab, mucus (3 points) Invasive:	0-5	2 (blood)

	blood (2 points) tissue and other (1 point) None (0 points)		
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) No (0 points)	0-3	3
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - please report it in the comment field, add the link (3 points) No (0 points)	0-3	3 AOP #296 (draft); Cho, E., Allemang, A., Audebert, M., Chauhan, V., Dertinger, S., Hendriks, G., Luijten, M., Marchetti, F., Minocherhomji, S., Pfuhler, S., Roberts, D. J., Trenz, K., & Yauk, C. L. (2022). AOP report: Development of an adverse outcome pathway for oxidative DNA damage leading to mutations and chromosomal aberrations. Environmental and molecular mutagenesis, 63(3), 118–134. https://doi.org/10.1002/em.22479
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	High - please provide proof of evidence via DOI or ref. in the comment field (5 points) Medium (3 points) Low (1 point) No adversity expected (0 points)	0-5	3 Costa, S., Carvalho, S., Costa, C., Coelho, P., Silva, S., Santos, L. S., Gaspar, J. F., Porto, B., Laffon, B., & Teixeira, J. P. (2015). Increased levels of chromosomal aberrations and DNA damage in a group of workers exposed to formaldehyde. Mutagenesis, 30(4), 463–473. https://doi.org/10.1093/mutage/gev00 2
Questions for assessing applicability (score 0-16)			
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Yes - please report it, provide the DOI in the comment field (5 points) Can be applied in a modified form (3 points) No (0 points)	0-5	Yes (5) Wultsch, G., Setayesh, T., Kundi, M., Kment, M., Nersesyan, A., Fenech, M., & Knasmüller, S. (2021). Induction of DNA damage as a consequence of occupational exposure to crystalline silica: A review and meta-analysis. Mutation research. Reviews in mutation research, 787, 108349. https://doi.org/10.1016/j.mrrev.2020. 108349

6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	Yes - please report it, provide the DOI (3 points) No (0 points)	0-3	3 Martins, M., & Costa, P. M. (2015). The comet assay in Environmental Risk Assessment of marine pollutants: applications, assets and handicaps of surveying genotoxicity in non-model organisms. Mutagenesis, 30(1), 89–106. https://doi.org/10.1093/mutage/geu037 Gajski, G., Žegura, B., Ladeira, C.,Pourrut, B., Del Bo', C., Novak, M.,Sramkova, M., Milić, M., Gutzkow, K. B., Costa, S., Dusinska, M., Brunborg, G., & Collins, A. (2019a). The comet assay in animal models: From bugs to whales - (Part 1 Invertebrates). Mutation research. Reviews in mutation research, 779,82–113.https://doi.org/10.1016/j.mrrev.2019.02.003 Gajski, G., Žegura, B., Ladeira,C.,Novak, M., Sramkova, M., Pourrut,B., Del Bo', C., Milić, M., Gutzkow, K.B., Costa, S., Dusinska, M., Brunborg,G., & Collins, A. (2019b). The comet assay in animal models: From bugs to whales - (Part 2 Vertebrates). Mutation research.Reviews in mutation research, 781, 130–164. Https://doi.org/10.1016/j.mrrev.2019.04.002
7) How would you define workload, training need related to an applicability for environmental settings?	High - minimal workload and training necessary (5 points) Medium - moderate workload and training necessary (3 points) Low - high workload and expert judgement necessary (0 points) Certified commercial labs offer this bioanalysis (add 3 points)	0-8	5
Questions for assessing validation and cost (score 0-16)			

8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes - publicly available in a peer-reviewed journal. Please provide DOI in the comment field (4 points) Partially - as internal or non-peer-reviewed SOP (2 points) No (0 points)	0-4	4 Collins, A., Møller, P., Gajski, G., Vodenková, S., Abdulwahed, A., Anderson, D., Bankoglu, E. E., Bonassi, S., Boutet-Robinet, E., Brunborg, G., Chao, C., Cooke, M. S., Costa, C., Costa, S., Dhawan, A., de Lapuente, J., Bo', C. D., Dubus, J., Dusinska, M., Duthie, S. J., Azqueta, A. (2023). Measuring DNA modifications with the comet assay: a compendium of protocols. Nature protocols, 18(3), 929–989. https://doi.org/10.1038/s41596-022-00754-y
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes - OECD guideline. Please provide links in the comment field (3 points) Yes - standardized DIN EN ISO. Please provide links in the comment field (3 points) No (0 points)	0-6	3 in vivo Comet assay was accepted to be included in the OECD Guidelines for the Testing of Chemicals (TG489)
10) What is the cost per sample?	Very low - <100 EURO/sample (6 points) Low - 100-250 EURO/sample (4 points) Medium - 250-400 EURO/sample (2 points) High - 400-750 EURO/sample (1 points) Very high - >750 EURO/sample (0 points)	0-6	6
Questions for assessing sensitivity & specificity & robustness (score 0-12)			
11) Is the Limit of Quantification (LOQ) below an accepted	Yes - please provide reference or DOI and	0-4	0

environmental exposure limit for a relevant reference substance?	LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)		
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4	0
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Yes - please provide reference or DOI and LOQ in the comment field (4 points) Partially (2 points) No or unknown (0 points)	0-4	2 Bonassi, S., Ceppi, M., Møller, P., Azqueta, A., Milić, M., Neri, M., Brunborg, G., Godschalk, R., Koppen, G., Langie, S. A. S., Teixeira, J. P.,Bruzzone, M., Da Silva, J., Benedetti, D., Cavallo, D., Ursini, C. L., Giovannelli, L., Moretti, S., Riso, P., Del Bo', C., hCOMET project (2021). DNA damage in circulating leukocytes measured with the comet assay may predict the risk of death. Scientific reports, 11(1), 16793. https://doi.org/10.1038/s41598-021-95976-7

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

B10: Characterisation of Genotoxicity measured via Micronucleus assay for reticulocytes (HH)

Name of the effect biomarker:

Micronucleus assay in Transferrin-positive Reticulocytes

Short description of the effect biomarker:

The frequency of micronucleated reticulocytes reflects chromosomal damage that has occurred in bone marrow approximately three days prior to blood sample collection. Analysis of micronuclei in young transferrin positive (+CD71) reticulocytes, that have just entered peripheral blood circulation, allows detection of systemic genotoxicity in humans exposed to radiation, genotoxic chemicals or any other conditions that cause DNA double strand breaks or disturb the mitotic apparatus. Flow cytometric analysis of enriched transferring positive reticulocytes enables rapid analysis of large amounts of cells, thus enhancing sensitivity and objectivity of the assay compared to microscopical analysis.

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

Tissue (Bone marrow), Organism (Systemic genotoxicity)

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Radioiodine therapy: 10.1016/j.mrgentox.2005.01.010 and 10.1016/j.toxlet.2004.12.007 (Ex vivo induction of reticulocyte micronuclei is not possible.)

Question No	Answer Options	Score
<u>Questions for assessing relevance and invasiveness (score 0-16)</u>		
1) Has the biomarker been assessed in easily accessible matrices?	blood (2 points)	2
2) Is there a plausible MoA?	Yes - please report your MoA in the comment field (3 points) Covers the endpoints of structural and numerical chromosomal aberrations caused by either clastogens or aneugens (https://doi.org/10.3389/fgene.2013.00131)	3
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - please report it in the comment field, add the link (3 points) AOP 296 Oxidative DNA damage leading to chromosomal aberrations and mutations)	3
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	Medium (3 points) Reticulocyte micronucleus assay allows detection of acute exposure that has occurred approximately three days prior to blood	3

	sampling. Lifespan of micronucleated reticulocyte is only some hours, thus, the method does not allow detection of longer-term cumulative effects. If the exposure is sufficient and sampling schedule has been correctly adjusted, biomarker is able to detect early biological effects of occupational exposure e.g., Hexavalent chromium: 10.3390/toxics10080483	
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in meaningful prediction/indication of ecological hazard?	Yes - please report it, provide the DOI in the comment field (5 points) Hexavalent chromium: 10.3390/toxics10080483 Pesticides: 10.1080/15287394.2011.582024	5
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	Yes - please report it, provide the DOI (3 points) Industrial pollution:10.1016/j.mrgentox.2020.503170 Smoking: 10.1096/fj.04-2729fje Strongly heated carbohydrate-rich food: 10.1016/j.fct.2018.08.029	3
7) How would you define workload, training need related to an applicability for environmental settings?	Medium - moderate workload and training necessary (3 points) Commercial magnetic bead isolation kit with extensive instructions available for the enrichment of transferrin positive reticulocytes. Sample staining and analysis principles can vary depending on the model of flow cytometry instrument. Certified commercial labs offer this bioanalysis (add 3 points) Only Litron laboratories, USA https://litronlabs.com/Products/In-Vivo-Human-Micronucleus	3+3
Questions for assessing validation and cost (score 0-16)		

8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes -publicly available in a peer-reviewed journal. Please provide DOI in the comment field (4 points) Method to isolate and analyze transferrin-positive reticulocytes: 10.1002/1098-2280(2000)36:1<22::aid-em4>3.0.co;2-u Single-laser flow-cytometry to measure micronuclei in an enriched transferrin-positive reticulocyte population: 10.1096/fj.04-2729fje	4
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes - OECD guideline. Please provide links in the comment field (3 points) OECD TG474 Mammalian Erythrocyte Micronucleus Assay https://doi.org/10.1787/9789264224292-en	3
10) What is the cost per sample?	Very low - <100 EURO/sample (6 points)	6
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	No or unknown (0 points)	0
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Partially (2 points) The assessment of MN RET is more sensitive in detecting biological effects from recent exposure. Due to the limited lifespan of reticulocytes, the suitability of the MN RET assay for the detection of low-dose chronic exposures in occupational settings must be studied further. (10.3390/toxics10080483) The bone marrow has to be affected to a high enough extent, and in the case of acute exposure optimal sampling time needs to be established in detail. (10.1016/j.toxlet.2004.12.007)	2
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Partially (2 points) The assessment of MN RET is more sensitive in detecting biological effects from recent exposure. Due to the limited lifespan of reticulocytes, the suitability of the MN RET assay for the detection of low-dose chronic	2

	exposures in occupational settings must be studied further. (10.3390/toxics10080483) The bone marrow has to be affected to a high enough extent, and in the case of acute exposure optimal sampling time needs to be established in detail. (10.1016/j.toxlet.2004.12.007)	

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

B11: Genotoxicity via Micronucleus Assay in mammalian cells for water quality assessment (DIN EN ISO 21427, OECD 487) (ENV)

Name of the effect biomarker:

Micronucleus Assay in mammalian cells for water quality assessment (DIN EN ISO 21427, OECD 487)

Short description of the effect biomarker:

The *in vitro*-micronucleus test is a genotoxicity test for the detection of micronuclei (MN) in the cytoplasm of interphase cells. Micronuclei may originate from acentric chromosome fragments (i.e., lacking a centromere), or whole chromosomes that are unable to migrate to the poles during the anaphase stage of cell division. The assay detects the activity of clastogenic and aneugenic test substances in cells that have undergone cell division during or after exposure to the test substance.

Biological level of complexity (Molecular or Biochemical, Cellular, Blood parameter, Tissue/Organ, Organism)?

MN assay for dividing cells (blood lymphocytes, tissue cells) can detect DNA damage at a cellular level. The primary DNA damage can result in a mutational cascade, finally leading to cancer.

Please add suitable reference compounds which can be used as positive control in dose-response relationships:

Methylmethanesulfonate (CAS 66-27-3) Mitomycin C (CAS 50-07-7) 4-Nitroquinoline-N- Oxide (CAS 56-57-5) Cytosine arabinoside (CAS 147-94-4) or Benzo(a)pyrene (CAS 50-32-8)

Question No	Answer Options	Score
Questions for assessing relevance and invasiveness (score 0-16)		
1) Has the biomarker been assessed in easily accessible matrices?	Non-invasive: water (5 points) Invasive: blood (2 points)	5 or 2
2) Is there a plausible MoA?	Yes – depends on the substance, which induces MN formation. MoA can be clastogenic or aneugenic finally leading to MN formation. see comment field	3
3) Is an Adverse Outcome Pathway (AOP) reported for this effect biomarker or is this effect biomarker situated in an AOP network?	Yes - depends on the substance, see comment field (3 points)	3
4) Is the biomarker able to detect relevant (adverse and severe) effects in environmental organisms during a long-term exposure?	Yes - See comment field	3
Questions for assessing applicability (score 0-16)		
5) Has the effect biomarker been applied in environmental studies and resulted in	Yes - please report it, provide the DOI in the comment field (5 points)	5

meaningful prediction/indication of ecological hazard?		
6) Has the biomarker been applied in risk assessment or other studies with regulatory relevance (e.g. discharge permitting, pesticide registration, contaminated site remediation)?	Yes - please report it, provide the DOI in the comment field (5 points)	3
7) How would you define workload, training need related to an applicability for environmental settings?	Certified commercial labs offer this bioanalysis (add 3 points)	3
Questions for assessing validation and cost (score 0-16)		
8) Does the biomarker have a well-described standard operating procedure (SOP)?	Yes - publicly available in a peer-reviewed journal. Please provide DOI in the comment field (4 points)	4
9) Does an OECD guideline or a standardized DIN EN ISO exist for the effect-biomarker?	Yes - OECD guideline (OECD 487). Test No. 487: In Vitro Mammalian Cell Micronucleus Test | en | OECD (3 points) Yes - standardized DIN EN ISO (DIN EN ISO 21427). DIN EN ISO 21427-2 - 2009-08 - Beuth.de (3 points)	6
10) What is the cost per sample?	High - 400-750 EURO/sample (1 points) Analysis costs depend on method and number of samples	1
Questions for assessing sensitivity & specificity & robustness (score 0-12)		
11) Is the Limit of Quantification (LOQ) below an accepted environmental exposure limit for a relevant reference substance?	No or unknown (0 points) There is no LOQ for genotoxic substances, just a risk estimation	0
12) Is the specificity* of the biomarker sufficient for the substances or effects of concern?	Partially (2 points) Specificity is lower than 80% Evaluation Example: Evaluation of the sensitivity and specificity of in vivo erythrocyte micronucleus and transgenic rodent gene mutation tests to detect rodent carcinogens - PubMed (nih.gov)	2
13+14) What are geometric mean concentrations and geometric standard deviations of the biomarker in the general population? Is the effect biomarker sufficiently robust to compare different levels of exposure risks (e.g. does it have age or seasonal variations or another dependency)?	Partially (2 points) - depends on the investigated cell type (e.g., lymphocytes or buccal cells) Yes, it is dependent on several factors (e.g., age, susceptibility).	2

* Specificity means the method is less prone to detect false positive results, e.g. specificity of 95% means only in 5 % of cases the biomarker detects false positive results. Weak point: We have currently no agreement which level of specificity is sufficient. It can be 90% or lower or higher. Preliminary we can work with 80% specificity as preliminary threshold.

Comments:

to 2.) **MoA (Mode of Action)**	yes, but it depends on the substance, which induces MN formation. MoA can be clastogenic or aneugenic finally leading to MN fomation. MN can originate during anaphase from lagging acentric chromosome or chromatid fragments caused by misrepair of DNA breaks or unrepaired DNA breaks. Malsegregation of whole chromosomes at anaphase may also lead to MN formation as a result of hypomethylation of repeat sequences in centromeric and pericentromeric DNA, defects in kinetochore proteins or assembly, dysfunctional spindle and defective anaphase checkpoint genes. https://doi.org/10.1093/mutage/geq052
to 3.) **Adverse Outcome Pathway (AOP) for this effect biomarker**	- depends on the substance, because an AOP is a model that identifies the sequence of molecular and cellular events required to produce a toxic effect when an organism is exposed to a substance - AOP can be applied to assess modes of action of genotoxins. Case studies are needed. doi: 10.1016/j.scitotenv.2018.02.015
to 4.) **MN in environmental organisms**	Studies are available, investigating MN formation in environmental organisms, relevant effects can be detected in dependence of environmental contamination **Fish (examples)**: DOI: 10.1007/398_2021_76 DOI: 10.1016/j.mrgentox.2021.503335 DOI: 10.1007/s11356-014-2537-0 DOI: 10.1007/s11356-016-7080-8 **Mussel (examples):** DOI: 10.1002/em.20646 https://doi.org/10.1016/0166-445X(95)00009-S **Frog (example):** DOI: 10.1038/s41598-018-26168-z **Crocodile (examples):** DOI: 10.3390/ani11113178 DOI: 10.1016/j.ecoenv.2016.10.035
To 5.) **Has the effect biomarker been applied in**	Yes, there are several studies available. Here are just some examples: DOI: 10.1093/mutage/geh040

environmental studies and resulted in meaningful prediction/indication of ecological hazard?	- DOI: 10.1080/15287394.2021.1881854 DOI: 10.1016/j.chemosphere.2016.03.001
To 6.) **Has the biomarker been applied in risk assessment or other studies with regulatory relevance?**	Yes, here are some examples DOI: 10.1007/s11356-015-5993-2 DOI: 10.1007/s11356-022-18767-1 DOI: 10.1080/01480545.2016.1209772
To 8.) **Does the biomarker have a well-described standard operating procedure (SOP)?**	- the guideline is: OECD 487 -published SOP: DOI: 10.3390/ijms21041534 -SOP for MN assay in fish cells: daimon_sop29_micronucleusassay_v1.1_dk.pdf (daimonproject.com) **BUT, you need dividing cells for this assay! The applicability of this assay for non-dividing cells is limited! (**DOI: 10.1093/mutage/get026)

Annex C. Relevant triggers for occupational effect biomarker use

In the "guiding principles to advance occupational mixture risk assessment with effect biomarkers" (OECD 2025), the advantages and limitations of using air monitoring, exposure biomonitoring and effect-biomonitoring are outlined in Table 1. To maximize the effectiveness of effect biomarker assessments, their advantages should be triggered based on both general criteria and MoA -specific considerations.

General trigger criteria for using effect biomarkers:

- Hazardous substance exposures with the potential to cause adverse effects in workers are likely to be present in the workplace (Carcinogenic, Mutagenic, Reproduction toxic (C MR) or equivalent concern, for example ED*, NT & DNT)
- Air monitoring or exposure biomonitoring cannot address worker risks sufficiently (e.g. due to substances with relevant dermal uptake or due to lack of assessment values (OEL*, OBL*) for expected exposures)
- In situations in which exposures to complex or unknown mixtures are likely and can cause adverse effects (examples are mentioned in the footnote).
- In work situations in which PPE or RMM are used to protect from complex or unknown hazardous substance it is required to check their efficiency in an integrated way.
- In all situations in which cumulative risk assessment is recommended by regulatory bodies (not mandatory).

In all these situations, an integrative effect assessment might quantify the effects of hazardous exposures, allowing an OHP to inform employers and workers on the need to implement additional RMMs*.

MoA specific trigger criteria for using effect biomarkers:

- **For assessing genotoxicity**:

  Carcinogenic substance exposures according to GHS* Cat. 1A and 1B and corresponding EU* and ACGIH* classes and IARC* Group 1 carcinogens and Group 2A or 2B (probable/possible human carcinogens).

- **For assessing reproductive toxicity**

  In situations where reproductive toxic exposures are expected, the following categories are defined according to GHS:

  Category 1A: Known human reproductive toxicants.

  Mixtures containing ≥ 0.1% or ≥ 0.3 % of such a substance.

  Category 1B: Presumed human reproductive toxicants - largely based on animal studies

Mixtures containing ≥ 0.1% or ≥ 0.3 % of such a substance.

Category 2: Suspected human reproductive toxicants - Evidence from animal and/or human studies is limited. Mixtures containing ≥ 0.1% or ≥ 3 % of such a substance.

TT-assessments in male workers should be performed if workplace substances are known to lower TT-levels.

- **For assessing ED according to EU ED Hazard class and statements**

  Currently active at EU level, but also under discussion at global level: Mixtures containing ED HH 1 or 2 substances, according to EUH380 or EUH381, May cause endocrine disruption in humans or are suspected of causing endocrine disruption in humans.

  TT -assessments in male workers should be performed if workplace substances are known to lower TT-levels.

- **For assessing NT & DNT**

  In situations in which substances are known to cause adverse neurotoxic (NT) or developmental neurotoxic (DNT) effects in humans or in valid animal studies. Additionally, e. g. according to the Regulation (EC) No 1272/2008, there is a classification for specific target organ toxicity after single or repeated exposure to the nervous system, i.e. NT . Existing classifications for developmental toxicity may also cover DNT concerns (since there is currently no specific classification for DNT). Further evidence of potential DNT concerns can be obtained from the results of the newly developed DNT *in vitro* battery (OECD 2023), mechanistic understanding, structural similarity (i.e. read-across assessments) to known DNT-inducing substances, findings from observational studies and/or scientific peer-reviewed open literature.

Additional evidence: Exposures to substances identified in screening programs using relevant New Approach Methodologies (NAMs); e.g. Endocrine Disruptor Screening Program of the US Environmental Protection Agency, DNT *in vitro* battery (DNT IVB) of the OECD, EFSA Genotoxicity Database, etc.).

References

- European Commission 2008. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (Text with EEA relevance)
- OECD 2023. Initial Recommendations on Evaluation of Data from the Developmental Neurotoxicity (DNT) In-Vitro Testing Battery. Series on Testing and Assessment. No. 377. https://www.oecd.org/en/publications/initial-recommendations-on-evaluation-of-data-from-the-developmental-neurotoxicity-dnt-in-vitro-testing-battery_91964ef3-en.html (last access on 10.12.2024)