

Molecular annotation of AOPs guides the development of the next generation mechanistic chemical safety assessment and new approach methods

Laura Aliisa Saarimäki¹, Jack Morikka¹, Alisa Pavel¹, Seela Korpilähde¹, Giusy del Giudice¹, Antonio Federico¹, Michele Fratello¹, Angela Serra¹, and Dario Greco^{1,2,}*

¹ Finnish Hub for Development and Validation of Integrated Approaches (FHAIVE), Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

² Institute of Biotechnology, University of Helsinki, Helsinki, Finland

*Corresponding author: dario.greco@tuni.fi

Abstract

Mechanistic toxicology has emerged as a powerful framework to inform on the safety of chemicals and guide the development of new safe-by-design compounds. Although toxicogenomics provides support towards mechanistic evaluation of chemical exposures, the implementation of toxicogenomics-based evidence in the regulatory setting is still hindered by uncertainties related to the analysis and interpretation of such data. Adverse Outcome Pathways (AOPs) are multi-scale models that link chemical exposures to adverse outcomes through causal cascades of key events (KEs). The use of mechanistic evidence through the AOP framework is actively promoted for the development of new approach methods (NAMs) and to reduce animal experimentation. However, in order to unleash the full potential of AOPs and build confidence into toxicogenomics, robust and unified associations between KEs and patterns of molecular alteration need to be established.

Here, we hypothesised that systematic curation of molecular events associated with KEs would enable the modelling of AOPs through gene-level data, creating the much-needed link between toxicogenomics and the systemic mechanisms depicted by the AOPs. This, in turn, introduces novel ways of benefitting from the AOP concept, including predictive models, read-across, and targeted assays, while also reducing the need for multiple testing strategies. Hence, we developed a multi-step strategy to annotate the AOPs relevant to human health risk assessment. We show that our framework successfully highlights relevant adverse outcomes for chemical exposures with strong *in vitro* and *in vivo* convergence, supporting chemical grouping and other data-driven approaches. Finally, we defined and experimentally validated a panel of robust AOP-derived *in vitro* biomarkers for pulmonary fibrosis.

Introduction

Mechanistic aspects of chemical exposures have been long exploited in the context of academic research, resulting in the emergence of toxicogenomics and systems toxicology as independent fields (1,2). Although the mechanistic insight gained through the technologies employed in academia has been valued as supporting evidence in the regulatory setting, its incorporation into the regulatory framework is to date hindered by concerns related to the robustness and reproducibility of such data and its analysis (3). At the same time, the growing need for faster, cheaper, and more ethical approaches for chemical safety assessment have made mechanistic toxicology central for clarifying aspects important to regulatory decision making. Furthermore, uncovering exposure related mechanistic properties is emerging as a fundamental approach for the design of new drugs and chemicals (4,5). Hence, multiple high-end research initiatives are underway to drive the shift from traditional animal-based assessment of apical toxicity endpoints towards *in vitro* and *in silico* approaches supported by mechanistic evidence (6–8).

Adverse Outcome Pathways (AOP) emerged as models to organise biological mechanisms into causally linked sequences of multi-scale events to support chemical risk assessment (9). AOPs have since expanded beyond the limits of toxicology, showing their applicability in organising mechanisms of disease progression and adverse health outcomes (10,11), and could even be applied to assess beneficial effects of therapies. The mechanisms depicted by AOPs comprise a sequence of events that progress from the molecular initiating event (MIE) towards an adverse outcome (AO) through intermediate steps, key events (KEs), with biological complexity increasing as the AOP progresses. Individual KEs are connected by key event relationships (KER), that verbally explain the causal link between the events and provide context for the pathway.

The AOP concept quickly attracted attention due to its potential in tackling one of the major challenges in the shift away from traditional toxicology: deciphering systemic and long-term outcomes of chemical exposures without the use of animal experiments. While significant efforts still need to be made towards this goal, AOPs encompass the means to systematically guide the integration of *in vitro*-based evidence into the risk assessment framework (12). AOPs provide the grounds for various predictive approaches, read-across, and the development of targeted assays and new approach methods (NAMs), as also suggested by regulatory agencies and international organisations, such as the OECD (8). Furthermore, the construction of AOPs can help identify gaps in knowledge and guide resources towards mechanisms in need of further investigation, or alternatively, reveal connections that have not been previously characterised (13).

Concurrently with the development of the AOP framework, the role of omics data in elucidating biomarkers and mechanisms of action of chemical exposures and diseases has become more prominent (14–18). Omics data have been used to support the development of AOPs, especially through the identification of molecular targets and mechanisms (19–23). However, full exploitation of omics-based evidence in the context of AOPs is hindered by the complication of linking molecular data to complex biological events, affecting both the development and the application of AOPs. Furthermore, while the value of omics data in answering questions of regulatory importance is recognised, the lack of standardisation in analysis and reporting have hampered widespread regulatory acceptance of omics-based evidence (24). Bypassing these challenges could broaden the application of AOPs and further aid in the development of the concept towards quantitative models and assays. While

molecular assays based on arbitrarily selected reporter genes have been proposed (e.g., ToxCast assays), there is an urgent need to develop new data-driven unbiased molecular assays for reliable and efficient mechanistic safety assessment of chemicals.

Here we hypothesised that rigorous curation of molecular events associated with AOPs could facilitate the implementation of omics-based evidence to 1) support the development of new AOPs, 2) identify and fill gaps in knowledge, and 3) transfer AOP-based knowledge into robust assays to support chemical safety assessment.

Well-curated gene ontologies, pathways, and biological processes are used to interpret omics results and their translation into biologically relevant information. While some KEs can be easily crosslinked with such terms and their associated genes, the annotation of complex KEs taking place at a higher level of biological organisation (e.g., at the tissue- or organism-level) is a more demanding task. This requires knowledge regarding ontologies and the biological events themselves. For instance, generic annotations are helpful for categorising KEs, but without the intention of modelling KEs using the associated gene sets, they will likely not reach the level of granularity required for such a task. This is currently reflected in the annotations provided in the AOP-Wiki repository (aopwiki.org). The annotation of KEs to selected ontologies is currently included as an option in AOP-Wiki. However, the coverage of the annotations is currently low and has not been intended for modelling the KEs using the gene sets associated with their annotations.

Previous efforts to curate external annotation have shown the potential of the approach (25,26). However, these have either remained at the level of abstract associations or focused on individual examples (27,28). Hence, systematic, fit-for-purpose, and up-to-date annotation linking KEs to curated gene sets has not yet been established. To this end, we applied an integrated strategy for defining gene-KE-AOP associations through systematic curation. We show the applicability of our strategy for evaluating potential adverse outcomes of chemical exposures, and for the identification of AOP-driven biomarkers that can inform the development of target assays and novel approaches to chemical hazard characterisation.

Results and discussion

We developed an integrated approach to systematically associate curated gene sets to the KEs and AOPs. Our approach combines natural language processing (NLP) techniques with manual curation to link relevant biological processes and pathways, as well as their associated genes, to KEs of AOPs relevant for human health risk assessment. The resulting gene-KE-AOP connections enable the modelling of KEs and AOPs through gene-level data, which further introduces novel ways to benefit from the AOP concept. We applied this approach to generate an AOP fingerprint for a known profibrotic exposure *in vivo* and *in vitro* and finally combined the annotation to a framework for prioritising KE- and AOP-associated genes to guide the discovery of biomarkers and reporter genes. The complete approach described in the following sections is summarised in Figure 1.

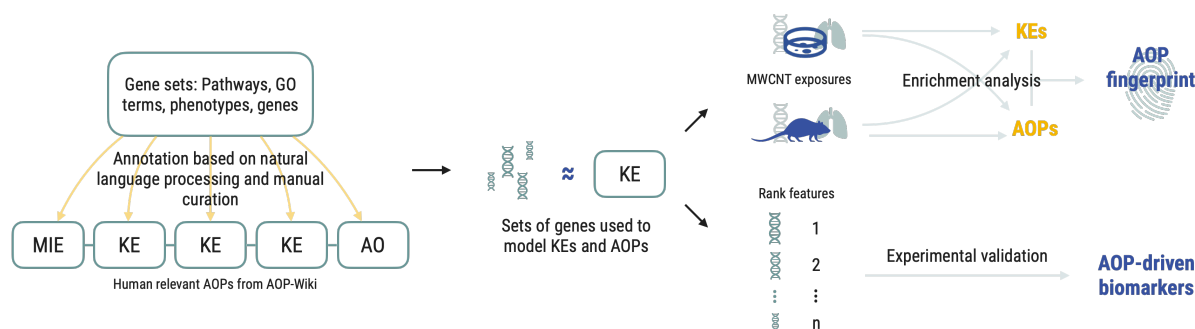


Figure 1. Overall scheme of the study. Established gene sets were annotated to KEs of the AOPs relevant for human health risk assessment. The resulting gene sets were then used to model the KEs. The validity of the annotation was evaluated using gene signatures of exposures with known adverse outcomes. Finally, we combined the approach with a gene prioritisation framework resulting in the identification of AOP-driven biomarkers for pulmonary fibrosis.

The majority of KEs can be successfully annotated to curated gene sets

At the time of retrieving the data from the AOP-Wiki repository (November 2020), a total of 289 AOPs and 1,131 distinct KEs were identified. However, after eliminating the AOPs for which taxonomic applicability was either not available nor in the scope of our investigation, 176 AOPs and 856 unique KEs remained, forming a total of 1,245 unique AOP-KE pairs (specific KEs). Although the AOP-Wiki houses selected annotations for some of the KEs, majority of them were considered not to be specific enough for our purpose (i.e., KEs describing the dysregulation of a specific gene annotated to terms such as “gene expression”). Additionally, as the existing annotations only cover a part of the KEs, we decided to consistently curate the annotation of all KEs. As a result, 799 unique KEs mapped to 175 AOPs received a curated annotation. The KEs were treated as individual instances, hence the same KE mapped to multiple AOPs was always annotated to the same term(s). A summary of the number of terms annotated to the KEs is presented in Figure 2A along with the proportions of the different term sources (Figure 2B). GO biological processes (GO_BP) represent most of the mapped annotations, followed by GO molecular functions (GO_MF) and Human Phenotype Ontology (HPO). Since up to five annotations were provided for the KEs, the final gene sets used from herein comprise the union of the genes mapped to each annotated term. This structure allowed improved specificity, while also providing the possibility to further refine the gene sets using the hierarchical order implemented where applicable. The size of the gene sets associated to each KE range from one to 5,990 genes, with a median value of 81 genes. Consequently, when AOPs are modelled by combining the gene sets associated to their KEs, the gene set sizes range from 15 to 5,992 with the median size being 752 genes.

In total, the annotations comprise 15,825 genes. While the majority of genes are annotated to less than 5 KEs (9,044 genes), 1,434 genes have more than 20 KEs associated to them, and 50 genes have more than 80 associated KEs (Figure 2C). Although these numbers can be affected by annotation bias, e.g., certain genes are better researched and annotated than others, they can also guide the selection of AOP-driven biomarkers when specificity is of importance.

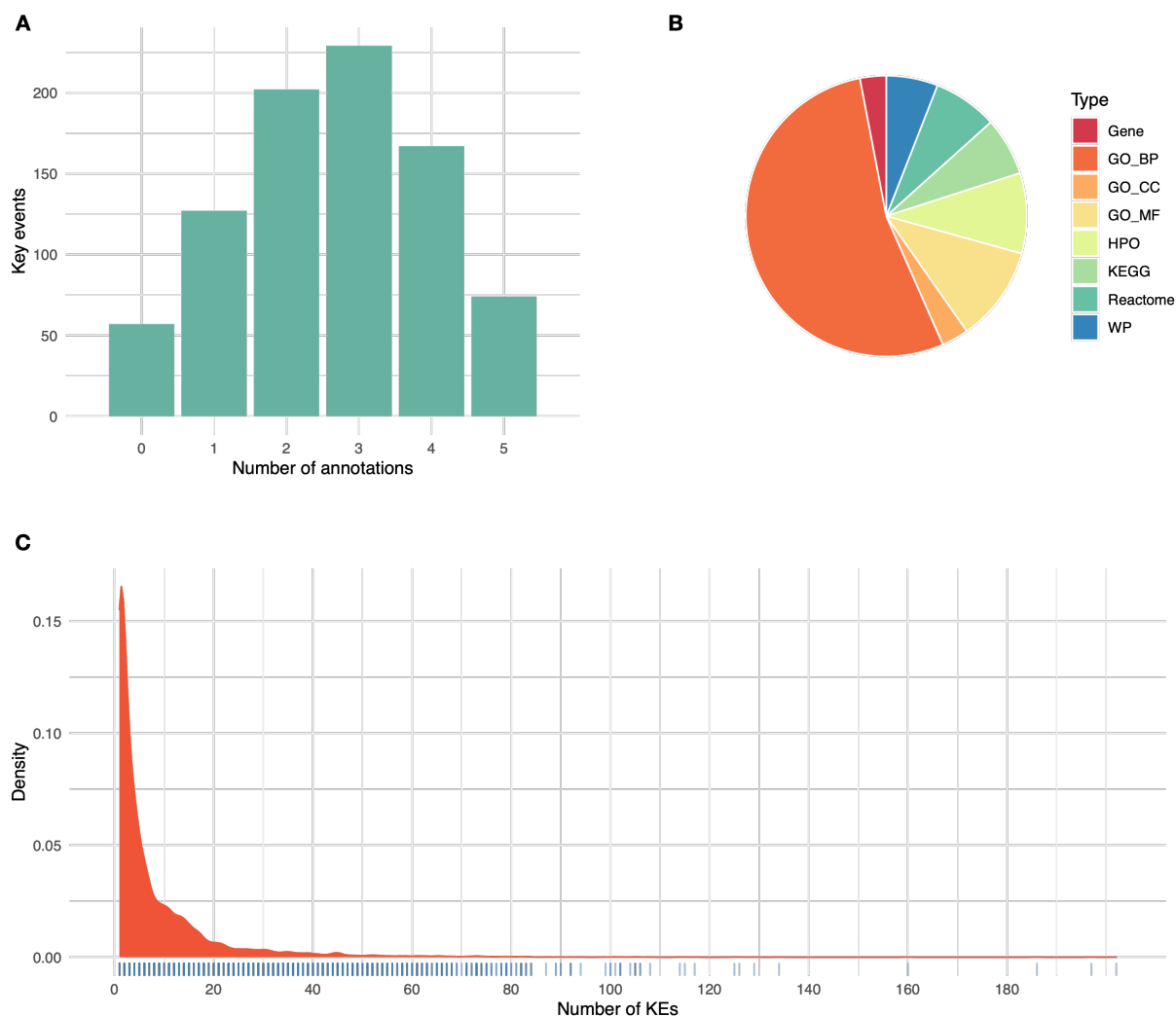


Figure 2: Descriptive analysis of the KE annotation. (A) Bar plot describing the number of annotated terms per KEs. (B) Pie chart expressing the proportions of different annotation types. (C) Density distribution of the number of KEs each gene is annotated to.

The result of the annotation suggests that the majority of human relevant KEs can be linked to gene ontology, phenotype, and pathway terms, and further to their corresponding sets of genes. Previous efforts to annotate KEs have provided similar implications (25,26). However, the applicability of the annotations has not been previously systematically explored, nor have they allowed modelling of the KEs and AOPs by use of gene sets on a large scale. Here, we present the first thorough mapping of all human relevant KEs to curated gene sets.

AOP enrichment highlights relevant adverse outcomes associated to chemicals

We tested the ability of our novel annotations to highlight relevant AOPs by analysing a set of curated reference chemicals as defined by EU Reference Laboratory for alternatives to animal testing (ECVAM) and National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). We focused on four categories of chemicals defined by their toxicity properties to include hepatotoxic and carcinogenic agents as well as thyroid disruptors and sex hormone receptor (estrogen receptor – ER, and

androgen receptor - AR) agonists. For each of the selected chemicals, we retrieved a list of associated genes from the Comparative Toxicogenomics Database (CTD) (29), resulting in a final set of 75 chemicals (Supplementary File 1).

AOPs related to each category were then identified among the 175 AOPs we had curated, and the prevalence of relevant AOPs (i.e., AOPs describing carcinogenesis for carcinogenic chemicals, etc.) among the five most significantly enriched AOPs for each chemical were evaluated. The results suggest that the enrichment approach successfully highlights AOPs of relevance for each group of chemicals (Figure 3). All sex hormone receptor agonists had at least one relevant AOP among the top five enriched, while the proportions varied from 43% (thyroid disrupters) to 93% (carcinogens) in the other categories (Figure 3).

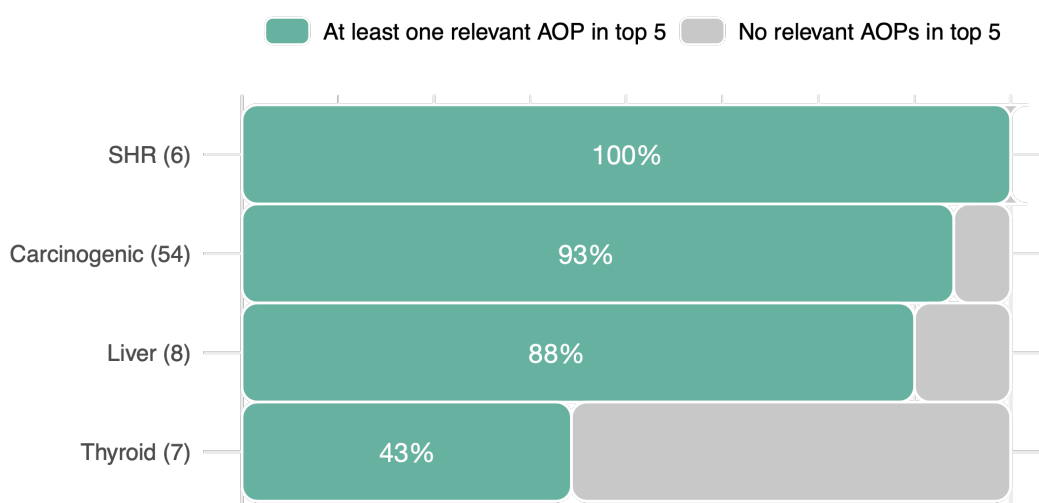


Figure 3: Bar plot representing the proportion of chemicals with relevant AOPs among the top five enriched AOPs based on the chemical classification. Number in brackets after the category name refers to the number chemicals in each category while the percentage on the bars reflects the proportion of chemicals in each category highlighting relevant AOPs. SHR stands for sex hormone receptor agonist.

In the group of carcinogenic chemicals, 93% of the compounds evaluated had cancer-related adverse outcomes among the top enriched AOPs. In fact, the group of carcinogens had the highest proportion of relevant AOPs at the top as compared to the others (median four out of five compared to the median of two out of five in the other groups). However, it should be noted that AOPs related to cancer are among the most represented group of AOPs, and cancer-related genes are generally highly researched and annotated, which may introduce a level of annotation bias that should be recognised.

The remaining four carcinogenic chemicals (7%) that showed no cancer AOPs among the top enriched AOPs were N-nitrosodiethanolamine, N-nitrosomorpholine, phenacetin, and tetrachloroethylene. N-nitrosomorpholine and N-nitrosodiethanolamine are both nitrosamines whose suspected adverse outcomes besides carcinogenesis include non-alcoholic steatohepatitis (30). Indeed, both compounds contained hepatic steatosis related AOPs among the top five enriched AOPs (Supplementary File 1). Tetrachloroethylene (perchloroethylene, PCE) is a chlorocarbon solvent used in dry-cleaning and other degreasing applications (31). AOPs with the most significant enrichment for PCE were also related to hepatic adverse outcomes. Although neurotoxicity is one of the most frequent AOs associated

with PCE exposure, hepatotoxicity has also been reported (31). Our results documenting liver steatosis are supported by biopsy-based evidence of liver disease, both in human and animal models, in settings of high occupational exposures (32). Lastly, phenacetin is a drug that was widely used as pain medication until it was withdrawn from the market across the globe due to increasing evidence of carcinogenicity and renal toxicity (33). The most enriched AOPs for phenacetin included immune related AOPs “Immune mediated hepatitis” (Aop:362) and Aop:277 titled “Inhibition of IL-1 binding to IL-1 receptor leading to increases susceptibility to infection”. Although there is no described association between phenacetin and IL-1 or immuno-toxicity, it is known that they both play a role in paracetamol-associated liver toxicity, which is the main metabolite of phenacetin (33,34).

In the case of the known liver toxicants, hexaconazole was the only compound not highlighting AOPs associated with liver toxicity among the top enriched AOPs. Hexaconazole is a widely used triazole fungicide. It acts by blocking sterol biosynthesis *via* inhibition of cytochrome P450 (35). Hexaconazole was considered as a Group C-Possible Human Carcinogen by the US EPA due to increased incidence of benign Leydig cell tumours in rats (https://www3.epa.gov/pesticides/chem_search/hhbp/R000356.pdf). Moreover, it was found to affect the reproduction of female rats (35). The top enriched AOPs correctly identify this signature. Furthermore, the top two pathways “HMG-CoA reductase inhibition leading to decreased fertility” and “Modulation of Adult Leydig Cell Function Subsequent to Decreased Cholesterol Synthesis or Transport in the Adult Leydig Cell” both suggest a decrease in cholesterol levels by inhibition of the HMG-CoA reductase. Drugs inhibiting this enzyme, such as statins, are known to possibly cause liver damage (36).

Known thyroid toxicants performed poorest in our analysis. Bifenthrin, malathion, permethrin and simazine did not capture thyroid-related AOPs among the top five enriched. All these compounds have been widely used in agriculture as herbicides or pesticides. Agrochemicals represent a significant class of endocrine disrupting chemicals, albeit through varying mechanisms. It is now accepted that many of these molecules may mimic the interaction of endogenous hormones with nuclear receptors, such as estrogen, androgen, and thyroid hormone receptors (37). Indeed, bifenthrin has already been reported as an endocrine-disrupting compound by blocking the binding of endogenous hormones (38). In our framework, its anti-estrogenic activity emerges as the most enriched AOP (Supplementary File 1). Malathion is an organophosphate pesticide that is known for its low acute toxicity and rapid degradation (39). In this light, it is not listed as a primary thyroid disrupting chemical, and its toxicity has been associated with the inhibition of acetylcholinesterase activity on nerve impulse (39). Recent studies, however, demonstrated that malathion acts as an endocrine disruptor, both *in vitro* and *in vivo* (40,41). Our results support these findings, highlighting the Aop:165: “Anti-estrogen activity leading to ovarian adenomas and granular cell tumors in the mouse” as well as Aop:112: “Increased dopaminergic activity leading to endometrial adenocarcinomas”. Furthermore, Moore *et al.* demonstrated that malathion exposure at higher concentrations induces cytotoxic and genotoxic effects in HepG2 through oxidative stress, which can finally lead to liver cancer (39). Similarly, our framework highlights both the “PPARalpha-dependent liver cancer” and “Cyp2E1 Activation Leading to Liver Cancer” AOPs. Simazine is a triazine herbicide whose use has been banned in most European countries for nearly two decades (42). Simazine has now been recognised, similarly to the other compounds, as an endocrine disrupter (42). Interestingly, the enrichment analysis for simazine highlighted AOPs related to the development of adenomas and carcinomas through endocrine disrupting activities (e.g., Aop:107 titled “Constitutive androstane receptor

activation leading to hepatocellular adenomas and carcinomas in the mouse and the rat”) as well as direct disruption of the GnRH pulse (Supplementary File 1). Although multiple *in vivo* and *in silico* evidence also indicate permethrin as possible endocrine disruptor (43,44), no endocrine related pathways are present in the top enriched AOPs. However, this framework was able to highlight the modulating effect of permethrin on the lipid metabolism. It has been demonstrated that in HepG2 cells, permethrin increases lipogenesis and decreases beta oxidation, possibly contributing to the development of NAFLD (45). Indeed, the “Inhibition of fatty acid beta oxidation leading to nonalcoholic steatohepatitis (NASH)” AOP is statistically enriched in our results.

Together, these results highlight relevant AOPs modelled by our curated gene sets to be enriched by the genes associated to the compounds, suggesting that our framework is able to support robust mechanistic and data-driven chemical grouping as well as the identification of potential adverse outcomes using chemical-gene associations.

Our annotation enables grouping of KEs resulting in improved modelling of the AOP network

In order to fully unleash the potential of mechanistic toxicology, more informative testing strategies need to be developed that can monitor specific phases of the exposure-bio interactions and mechanisms. To this end, we defined accurate sets of genes capable of modelling specific KEs and AOPs. However, one of the challenges observed in the AOP-Wiki is the redundant semantics in the naming of KEs. While creating a new KE can be meaningful in many cases (e.g., the same biological process taking place in a distinct organ or a tissue), unnecessary redundancy can lead to challenges in the application of the AOP-based knowledge. This is especially true when modelling AOPs as a network and using such representation to identify hidden connections and to perform read across analysis (10,46–51).

Hence, we hypothesised that KEs could be grouped based on the degree of similarity of their associated gene lists. We calculated the similarity of the KEs based on their annotated gene sets and grouped together those mapped to identical sets of genes (Jaccard Index = 1). This resulted in the identification of 637 groups of varying sizes. These groups were characterised by four main concepts: 1) truly duplicated KEs due to distinct semantics, 2) same biological event in multiple biological systems, 3) subsequent KEs mapped to the same terms due to inadequate specificity, and 4) opposite regulation of the same biological event (i.e., increased vs. decreased signalling).

Here, the grouping based on identical gene sets was selected due to the nature of the downstream application and statistical considerations (i.e., to avoid multiple testing against the same gene set in enrichment analysis). However, a parallel approach with varying cut-off values for similarity could be implemented to cluster KEs more roughly and to define specific categories of events. Similarly, further refinement of the KE clusters could help to enhance the AOP network by removing redundant nodes which, in turn, could reveal hidden links. The potential of the KE grouping was showcased using a subgraph formed by considering the AOPs related to pulmonary fibrosis (PF). PF is a chronic lung disease characterised by tissue damage and scarring that impairs lung function (52). A range of environmental exposures, including certain chemicals, drugs, nanomaterials and radiation, as well as infectious diseases have been identified as causative agents for PF (52–54). Moreover, the COVID-19 pandemic has raised concerns about increasing rates of PF (55–57). Understanding the disease

mechanisms can help in the development of strategies to treat and prevent the disease, and to control and modulate the exposures that contribute to its progression. Furthermore, it can serve as the foundation for developing targeted assays for evaluating PF as a toxicological endpoint.

Six AOPs related to PF were available in the AOP-Wiki at the time of data retrieval (Figure 4A). These distinct AOPs characterise multiple pathways leading to the same adverse outcome. Together, these AOPs comprise 30 KEs, which form a connected graph when modelled as a network (Figure 4C). However, several redundancies were observed among the KEs. For instance, the adverse outcome was expressed either as *lung fibrosis* (Event:1276) or *pulmonary fibrosis* (Event:1458). Hence, the application of the similarity-based grouping resulted in 23 distinct KEs (Figure 4B) that were then used as the basis for merging the KE nodes in the PF network (Figure 4D).

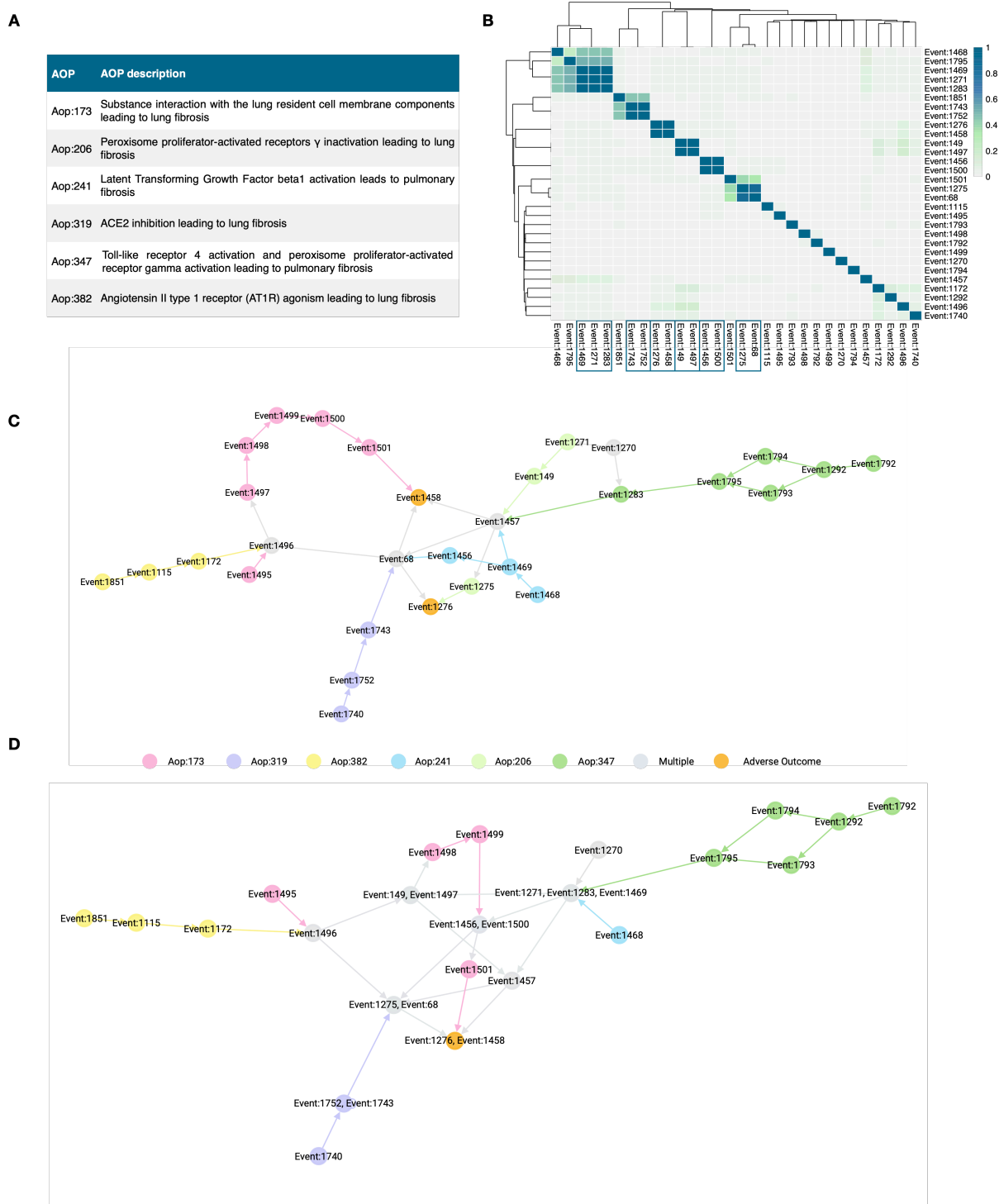


Figure 4: A) Table presentation of pulmonary fibrosis (PF) AOPs identified in the current study. B) Heatmap representing the Jaccard index-based similarity of the PF KEs as per their associated gene sets. Values close to zero (light grey) correspond to a low similarity between distinct KEs, while the increasing similarity is expressed with the colour changing through green to blue. C) Graph presentation of the PF AOPs using their original KEs. Distinct colours denote the KEs of individual AOPs, grey nodes are KEs shared by multiple AOPs, and orange nodes correspond to the shared adverse outcomes. D) Graph presentation of the PF AOPs after KE grouping. The number of shared (grey) nodes has now increased, and the duplicated AO has been grouped into one distinct AO (orange).

The PF AOPs formed a connected network, indicating that each of the individual AOPs shared at least one KE with one or more of the other AOPs. However, as the duplicated KEs were merged, the similarities between the AOPs became more evident. This is evidenced by the increasing number of shared KEs in the graph after merging (the grey nodes in Figure 4D) as compared to the original graph (Figure 4C). Furthermore, the merging revealed Aop:206 to be fully contained within the other AOPs.

The refinement of the AOP network through KE grouping simplifies the network while also enhancing the robustness of the KE relationships, depicted by the connections between the nodes. This process, in fact, removes redundant nodes, which supports the application of AOP networks in AOP research and risk assessment. Furthermore, as duplicated events are removed, the true influence of each node can be assessed more robustly through network analytics.

This example demonstrates the effect of KE redundancy and the potential of data-driven grouping of the KEs. While manual assessment and grouping would be achievable for a limited number of AOPs at a time, doing it AOP-Wiki wide would be a massive undertaking. Here, we show how our curated gene-KE-AOP connections can help guide the grouping and hence enhance network-based approaches in AOP research. Furthermore, our results suggest that it is often possible to identify meaningful genes that can successfully monitor multiple key events of a similar nature.

The AOP fingerprint of multi-walled carbon nanotubes converges *in vitro* and *in vivo*

Toxicogenomics has supported the development of mechanistic toxicology and further enhanced the possibility to obtain relevant information from *in vitro* studies, which could reduce the need for animal experimentation (58–60). Here, we tested the hypothesis that toxicogenomic data generated in two independent *in vitro* and *in vivo* exposure models would converge on a robust set of relevant AOPs. We focused on Mitsui-7, a known hazardous long and rigid multi-walled carbon nanotube (MWCNT). The airways provide the most prominent route of exposure to this nanomaterial, and it is best characterised for its lung-related adverse outcomes, including fibrosis (61–64). Hence, we selected data derived from a lung exposure to the MWCNT in mice (65), and an *in vitro* dataset with exposures on four cell lines representative of the human airways (59,66). These cell lines include differentiated THP-1 cells as a model of macrophages, A549 representing alveolar basal epithelial cells, BEAS-2B as bronchial epithelial cells, and MRC as a model of lung fibroblasts. Differentially expressed genes (DEGs) from all experimental conditions *in vivo* and each cell line *in vitro* were obtained from Saarimäki et al. (67) and merged into a single mechanism of action *in vivo* and *in vitro*, respectively.

We then performed enrichment analysis against both the AOPs and the KEs separately in order to evaluate the coverage of distinct KEs. We used the proportion of significantly enriched KEs to further filter the significant AOPs. This led us to identify 33 significant AOPs from the *in vivo* data, while 12 resulted significant from the *in vitro* exposure. These results were defined as the specific AOP fingerprint for the exposures, and it is presented in Figure 5.



Figure 5: AOP fingerprint of Mitsui-7 exposure *in vitro* and *in vivo*. Size of the bubble reflects the proportion of significantly enriched KEs in an AOP, while to colour denotes the FDR-adjusted p-value in a negative logarithmic scale (i.e., the higher the number, the smaller the p-value). The AOPs are organised by the enrichment p-value from the *in vivo* data.

Despite the distinct sizes of the AOP fingerprints, 10 of the 12 AOPs enriched *in vitro* were also included in the *in vivo* fingerprint. Moreover, the top enriched AOPs were shared and ranked similarly between *in vivo* and *in vitro* when ranked by the smallest adjusted p-value. The AOP enriched with the most significant p-value in both instances was Aop:173 titled “Substance interaction with the lung resident cell membrane components leading to lung fibrosis” (Figure 5). The *in vivo* data set was able to capture 7 of the 8 KEs of the AOP as significantly enriched, while 3 out of the 8 KEs were enriched *in vitro*. Interestingly, Aop:173 has been specifically developed with robust evidence from MWCNT exposures, and multiple types of carbon nanotubes are listed as known stressors for the AOP (<https://aopwiki.org/aops/173>). The second AOP (Aop:171), on the other hand, describes the induction of pleural/peritoneal mesotheliomas by chronic cytotoxicity in rats. Like most AOPs used in this study, Aop:171 is still under development and lacks information on potential stressors. However, mesothelioma is a well-known adverse outcome of asbestos exposure, a fibrous silicate mineral whose adverse effects have often been used as a warning example for MWCNTs (68). Indeed, similarities in their mechanism of action have been extensively investigated (61,69,70).

The *in vitro* AOP fingerprint captures effects such as frustrated phagocytosis, oxidative stress, cytotoxicity, and immune activation, which have all been reported as consequences of this type of exposure and contribute to the pathogenic nature of Mitsui-7 (61,62,64,71).

Similarly, the profibrotic effects are highlighted with the multiple PF AOPs enriched. These effects are also observed in the *in vivo* AOP fingerprint. However, the *in vivo* fingerprint further highlights various AOPs outside the respiratory system, which is less apparent *in vitro*. While adverse outcomes beyond the immediate exposure site are feasible, many of these could likely be accounted for by the different effects of similar transcriptomic signatures in different biological systems (e.g., multiple AOPs related to gastric ulcer formation could reflect similar mechanisms of surfactant disturbance in two distinct exposure sites). On the other hand, the AOPs unique to the *in vitro* fingerprint, Aop:277 and Aop:263 (Figure 5), reflect the specific effects of the Mitsui-7 exposure on the immune system. Such specific signals can be easily masked in the *in vivo* system, where a large array of cell types is affected by the exposure.

It is worth noting that the exposures selected for the analysis had diverse set ups and a notable difference in the size of the combined mechanism of action (863 DEGs *in vitro* vs. 3,540 *in vivo*). While data from multiple cell lines were selected to capture effects besides immune cell activation *in vitro*, we were not able to match the dose and time course set up present in the *in vivo* dataset. However, we wanted to include this long-term exposure to evaluate whether it would result in broader coverage over the KEs of AOPs. Furthermore, histopathological evaluation from the same *in vivo* exposure set up has shown fibrosis in the lung from the day 7 onwards (72), suggesting that a whole PF AOP could be covered with this data. Indeed, all but the MIE (Event:1495) of Aop:173 were enriched *in vivo*. The high proportion of enriched KEs in the *in vivo* data supports the modelling of KEs with relevant gene sets and the use of toxicogenomic evidence for the development of AOPs, as well as the evaluation of potential adverse outcomes of chemical exposures. Likewise, we show that the analysis of toxicogenomic data against robustly annotated AOP framework supports a high degree of *in vitro* to *in vivo* extrapolation and further supports the inclusion of toxicogenomics-based evidence for regulatory purposes.

KE-associated gene sets guide the selection of biomarkers

We showed that our KE-linked gene sets provide a robust way of evaluating potential outcomes of chemical exposures from transcriptomics data. This observation alone can help to guide chemical testing and grouping. However, to support the development of target assays and integrated approaches, specific reporter genes and markers need to be identified.

Selection of transcriptional biomarkers and reporter genes only based on differential expression from experimental data gives little context or reference to the AO of interest. Even if a certain exposure is known to induce a specific endpoint, there is no indication whether the measured deregulation could be associated with the phenotype of interest. On the other hand, prioritising the features in the context of the KEs or whole AOPs could shed light on the importance and specificity of the feature regarding the phenotype. This, in turn, can guide the selection of potential biomarkers even in the absence of experimental data. Hence, we implemented a universal and customisable framework for the prioritisation of the KE-associated genes to drive the identification of AOP-informed biomarkers and used it to identify AOP-driven biomarkers for PF. The shortlisted marker genes were then screened by RT-qPCR in an *in vitro* model of human macrophages exposed to bleomycin, a well-known profibrotic chemical (73).

First, we defined characteristics for optimal biomarkers based on the Bradford Hill criteria, originally defined to evaluate causality in epidemiological research (74), but later

adopted to other research fields as well (75). Our newly defined characteristics, their Bradford Hill counterparts, and short descriptions of the consideration of each step in the selection process are summarised in Table 1.

Table 1: We defined characteristics for optimal biomarkers based on the Bradford Hill criteria. The characteristics were then implemented into the prioritisation and selection protocol, and further to the evaluation of the prioritised genes.

Bradford Hill	Our characteristic	Method/Assessment
Consistency (reproducibility)	Reproducibility	Selection considers evidence from previous profibrotic exposures
Strength (effect size)	Amplitude	Significant alteration of the expression as compared to control
Experiment	Measurable	Transcriptional biomarkers measurable by qPCR; selected genes need to be expressed in the model
Biological gradient (dose-response relationship)	Dose-responsive	Benchmark-dose modelling to evaluate dose-response
Coherence	<i>In vitro</i> to <i>in vivo</i> extrapolation	Experimental evidence from <i>in vitro</i> and <i>in vivo</i> *
Analogy	Predictive (of the outcome of interest)	Selection based on the KEs preceding the AO of interest
Specificity	Specificity	Gene ranking based on the specificity score
Plausibility	(Biological) plausibility	The AOP framework provides a plausible context; supporting evidence; selection of the organism
Temporality	Temporality	Transcriptional alteration follows the exposure; selection of the model organism **
–	GLP-method	RT-qPCR
–	Influence	Centrality measures from human protein-protein interaction and gene regulatory networks

*The biomarkers selected here are targeted for the development of non-animal assays for toxicological assessment. Hence the coherence to *in vivo* set ups is not evaluated experimentally. However, *in vivo* data was used for the selection of the markers to provide context of the systemic response. **Temporality in the Bradford Hill criteria refers to a clear distinction of the exposure happening prior to the outcome. Here, we considered temporality by observing transcriptional changes post exposure as well as in the selection of the model organism. Macrophages have a crucial role in the initiation of the profibrotic response preceding the outcome, fibrosis.

The prioritisation and selection of the candidate biomarkers considered three main aspects: 1) the social life of genes, i.e., some genes (gene products) are more influential than others, 2) specificity regarding the endpoint of interest, and 3) experimental evidence suggesting the genes respond to a relevant exposure. The ranking of the genes was based on the first two considerations, while the experimental evidence was included to guide the selection of candidate genes for RT-qPCR validation from the ranked list. This enabled a flexible selection process that would be applicable even in the absence of experimental data. At this stage, we

also considered the biological feasibility of the target genes given the selected macrophage model as well as a broad coverage over the PF KEs.

As a result, we obtained a list of 25 candidates out of the original 2,075 genes related to PF (Table 2). Although we focused on the genes in the top 10%, we further included genes ranking lower to obtain a broader coverage over the PF KEs. Genes that are specific to individual KEs might rank low when the individual lists are combined. Hence, we considered the expression patterns from the experimental data as well as the specificity scores and ranks in the individual KEs. This also allowed us to evaluate whether genes ranked higher would perform better than others.

Table 2: Genes selected for qPCR validation. Green = yes, white = no.

Gene (rank)		Detected (Amplification)	Deregulated (ANOVA)	Dose-dependent
SMAD7 (1)	6h			
	24h			
	72h			
CXCL2 (3)	6h			
	24h			
	72h			
SPP1 (18)	6h			
	24h			
	72h			
CCL2 (19)	6h			
	24h			
	72h			
TGFB1 (23)	6h			
	24h			
	72h			
IL8 (33)	6h			
	24h			
	72h			
LOX (48)	6h			
	24h			
	72h			
PLOD2 (74)	6h			
	24h			
	72h			
MMP7 (80)	6h			
	24h			
	72h			
CXCL10 (91)	6h			
	24h			
	72h			
CCL7 (93)	6h			
	24h			
	72h			
MMP9 (105)	6h			
	24h			
	72h			
LTBP4 (112)	6h			
	24h			
	72h			
FN1 (116)	6h			
	24h			
	72h			
GDF15 (153)	6h			
	24h			
	72h			
MMP19 (179)	6h			
	24h			
	72h			
PTX3 (220)	6h			
	24h			
	72h			
TGFB3 (297)	6h			
	24h			
	72h			
	6h			

LTBP3 (335)	24h			
	72h			
TWIST1 (609)	6h			
	24h			
TGFBI (727)	72h			
	6h			
CTSK (759)	24h			
	72h			
RCN3 (1592)	6h			
	24h			
RSAD2 (1596)	72h			
	6h			
PLK3 (2027)	24h			
	72h			

We could detect the expression of 22 of the candidate genes at one or more of the evaluated time points, and six of the detected genes showed significant alteration as compared to the unexposed control samples (Table 2). Finally, five of these genes were altered in a dose-dependent manner: CXCL2 and CCL7 at 24h, IL8 (CXCL8) at 24h and 72h, and MMP19 and TGFBI at 72h. All but TGFBI of these genes were among the top 10% in the global PF rank (Table 2). Although we were not able to fit a dose-dependent model on the highest ranked gene, SMAD7, a suggestive trend could be appreciated in its expression pattern (Figure S4 panel 18/6H). The expression of each of these genes was upregulated as compared to the controls (Figure S4).

The central role of TGF-beta signalling is well-established in PF (76), but neither of the TGFβ genes tested (TGFB1 and TGFB3) showed significant change in expression in our setup. Indeed, TGF-beta is activated through a complex cascade of events, where the inactive form of the protein is activated by other effectors post-translationally (77), making members of the TGF-beta family a more attractive target for protein-based biomarker assessment over gene expression. At the same time, we did observe upregulation of SMAD7 and TGFBI which are both activated by TGF-beta (78,79), suggesting the induction of TGF-beta signalling. The protein encoded by TGFBI is involved in the extracellular matrix (ECM), and it has been shown to bind type I collagen, resulting in thicker fibres and further affecting macrophage polarisation towards the M2 type (80). Indeed, bleomycin has been suggested to polarise macrophages towards M2 (often referred to as the anti-inflammatory type), which have been shown to drive the development of PF through their ability to promote myofibroblast differentiation (81,82). Many of our suggested biomarkers are chemokines that mediate immune responses. IL8 and CXCL2 are best characterised as neutrophil attractants, while CCL7 targets a wide variety of leukocytes (83–85). Indeed, prolonged inflammation, combined with persistent M2 macrophage activation, supports pathogenesis of fibrosis (86), and a mixed status of M1/M2 macrophage activation has been previously associated with carbon nanotube -induced PF *in vivo* (87). Similarly, multi-walled carbon nanotubes have been shown to induce the polarisation of macrophages towards such mixed status of M1/M2 polarisation (88,89). MMP19 is a member of the matrix metalloproteinase family involved in ECM remodelling (90). MMPs have been extensively characterised in the context of PF (91,92), and MMP19 specifically has been suggested as a key regulator of PF in mice and humans (93).

Although macrophages alone cannot capture all the KEs of PF, our model is able to highlight the key steps of macrophage involvement in PF. The temporality of the expression

of our suggested biomarkers is supportive of the events leading to the development of fibrosis, where the initial inflammation is followed by type M2 macrophage activation that together contribute to the development of a profibrotic microenvironment and responses in other cells in the tissue (86).

New approach methods (NAMs) are urgently needed to reduce animal testing while providing robust evidence to support chemical safety assessment. While alternative methods have been successfully developed to capture acute effects, modelling long-term outcomes of the exposures, such as fibrosis, *in vitro* is still a challenge. Here, we propose a panel of five genes CXCL2, CCL7, IL8, MMP19, and TGFBI as AOP-derived robust biomarkers of PF to be successfully measured in a model of human macrophages *in vitro* after short exposure time.

Conclusions

Mechanistic toxicology encompasses the means for faster, cheaper, and more ethical chemical safety assessment. However, to unleash the full potential of mechanistic evidence also in the regulatory framework, robust approaches to build confidence towards toxicogenomics are urgently needed. Here, we introduced an integrated approach that links toxicogenomics with the concept of AOPs and proved its applicability to chemical grouping and development of data-driven NAMs. We introduce the concept of AOP fingerprint of a chemical exposure to evaluate potential systemic outcomes through unbiased interpretation of toxicogenomics data. Our results point to a consistent AOP fingerprint of multi-walled carbon nanotubes extrapolated from both *in vitro* and *in vivo* experiments. Finally, we identified and experimentally validated a panel of robust AOP-derived *in vitro* biomarkers for pulmonary fibrosis.

Materials and Methods

Definition of a Knowledge Graph -based data structure

We established a Knowledge Graph -based data structure by expanding our previously introduced framework, the Unified Knowledge Space (UKS) (94,95). A detailed description and a full list of integrated data sources are provided in the Supporting Methods. The so formed data structure was utilised throughout the study as described in the following sections.

Annotation of key events

We applied a multi-step strategy comprising natural language processing (NLP) and manual curation to annotate KEs to established gene sets through pathways, phenotypes, and gene ontologies. The annotation strategy is summarised in Figure S2.

Computational prioritisation of KE annotations: To match the descriptions of key events and gene sets we developed an NLP pipeline (Figure S2). The pipeline performs several operations to extract the informative terms from both the descriptions of a key event and a gene set that are scored to assess the degree of matching between the two entities. In detail, first, the raw text was converted to lower case and all punctuation symbols were removed. Second, concepts that span multiple words in the text description were replaced by a single word

expressing the same concept to strengthen the matching quality (e.g., the concept “positive regulation” was replaced with the single word “upregulated”). Third, the text was split into tokens which are further processed one by one. Fourth, each token corresponding to a stop word in the English language was dropped. Stop words refer to the most common words in a language that do not bring additional meaning (e.g., for the English language common stop words include “in”, “the”, “of”, “from”). Fifth, different declinations of the same concept were mapped to their root term (e.g., plurals were converted to singulars, the terms “increased” and “increasing” were both mapped to “increase”). We also used this same procedure to standardise several styles to write the same symbol (e.g., “ $\text{ppar}\alpha$ ” and “ $\text{ppar}\alpha$ ” map both to “ $\text{ppar}\text{-}\alpha$ ”). After these preprocessing steps, each gene set and key event was represented by a set of token words, e.g. {upregulate, $\text{ppar}\text{-}\alpha$ }. However, the frequency of each token word across the descriptions of genes and key events is not the same, and hence, the informative value of rare terms is higher than the informative value of more common tokens. We took this into account by weighting each token by its inverse document frequency (IDF), i.e., the weight was inversely proportional to the number of gene sets and key events that contain that token. Finally, we employed a weighted version of the Jaccard Index (JIW) to match gene sets and the key events, using the IDF as weights (i.e., each token that was shared between a gene and a key event did not account 1 as in the standard Jaccard Index, but it contributed its IDF weight to the matching score) and sorted the matching gene sets for each key event in descending order.

Manual curation and refinement of annotations: Next, the results of the computational prioritisation were manually evaluated for correct context and accuracy. Manual curation was used for gap filling and refinement of the annotations. In detail, the top five matches retained from the natural language processing -based approach were evaluated, and inaccurate or spurious matches were discarded. In case no matches from the computational prioritisation were deemed suitable, a manual search related to the name of the KE was performed on relevant databases (WikiPathways (96), HPO (97), KEGG (98), Reactome (99) and GO (100)). For molecular level KEs, where the alteration of an individual gene was described, either the main functions of the gene were selected, or the gene was directly annotated to the ensemble identifier of the said gene. More generic annotations (i.e., annotation of a KE describing the alteration of a gene to a functional term tightly related to that gene) were prioritised to increase the size of the relevant gene sets.

The matches for KEs expressing the increase/activation or decrease/repression of a biological process were further organised based on the hierarchy of the terms by prioritising the most generic but suitable term followed by increasing specificity when multiple annotations of various specificities were available. For instance, Ke:1457 called “Induction, Epithelial Mesenchymal Transition” was annotated to the following terms: 1) Epithelial to mesenchymal transition (GO:0001837), 2) Regulation of epithelial to mesenchymal transition (GO:0010717) and 3) Positive regulation of epithelial to mesenchymal transition (GO:0010718). The curated KE – gene set links were added to the UKS so that for each key event entity its top five matches were added, while the matching level was stored as an edge attribute. This allows to either combine multiple mappings for a key event or to filter for specific mapping levels. Since the KE – gene set mappings are always the same for the same KE, these relationships were added to the *Key Event* entities and not to the *Specific Key Event* entities, which reduces complexity of the knowledge graph as well as reduces needed storage space. The information, however, can still be retrieved from the UKS via its connecting paths.

Gene set retrieval: The genes corresponding to the matched terms were retrieved by matching the term names to their exact identifiers and querying the UKS for human genes associated with the terms. For phenotypes (HPO and KEGG disease), only genes with a link in the original database were included by filtering by the source for the connection. In cases where no human genes were linked to the annotated GO term, we obtained the mouse and rat genes associated and converted them to human orthologs using Ensembl (101), which were then used as the corresponding gene sets. When no genes of human, mouse, or rat were associated with the original term, the annotation match was discarded and considered unsuccessful. Once gene sets to all original terms were defined, the gene sets were merged to obtain the final set of genes corresponding to each KE in this study.

Enrichment analysis of reference chemical -associated gene sets

To evaluate the ability of our framework to highlight relevant adverse outcomes from chemical associated gene signatures, we retrieved lists of reference chemicals from the EU Reference Laboratory for alternatives to animal testing (ECVAM) reference chemical library (102) and National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) website (<https://ntp.niehs.nih.gov/whatwestudy/niceatm/resources-for-test-method-developers/refchem/index.html>). From the resources provided by ECVAM, we selected a hepatotoxic chemical list that had clear distinctions between positive and negative compounds. This list was based on the work from EPA's Virtual Liver project (https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=166616&Lab=NCCT), and was provided as a downloadable Excel-file (./CHELIST/CheLIST__EPA_VLIVER.xlsx) by ECVAM. From NICEATM, we selected the list of chemicals with characterised thyroid activity (specified as "ACTIVE" in the listing produced based on a previous publication by Wegner et al. (103). Androgen receptor and estrogen receptor agonists were selected from the lists of *in vitro* reference chemicals provided on the website. These listings had been previously published in Kleinstreuer *et al.* (104) and Browne *et al.* (105), respectively. Finally, carcinogenic compounds were identified from the list containing chemicals that are either known carcinogens or reasonably anticipated to be human carcinogens (RAHC) based on the 14th report on Carcinogens (RoC Classifications) provided by NICEATM. The chemicals from each of the reference lists were then matched to the list of chemicals obtained from the Comparative Toxicogenomics Database (CTD) (29) through name-based matching or by the provided CAS identifiers, resulting in the final lists of reference chemicals for each endpoint used in this study. Chemical-gene links originating from the CTD were retrieved from the UKS and only chemicals with 50-1,000 associated genes were included in the enrichment analysis.

Enrichment analysis was performed using the Fisher's exact test as implemented in the function *enrich* from R package bc3net (106) for each chemical associated gene set against the list of AOP-related genes (*i.e.*, the union of all the genes associated to all the KEs of the AOP). Enrichment p-values were adjusted using the false discovery rate (FDR) correction. AOP was considered significantly enriched with FDR-corrected p-value < 0.01.

KE clustering and construction of the pulmonary fibrosis network

Similarities between the gene sets associated to each KE were evaluated by calculating the Jaccard Index (JI) between all pairs of KEs (size of the intersection divided by the size of the union of the gene sets). The resulting similarity matrix was then used to group the KEs using hierarchical clustering as implemented in the function *hclust* in R package stats, specifying the

agglomeration method as “ward”. The number of clusters was defined so that only KEs with the same gene sets associated to them ($Jl = 1$) were assigned to the same group. The grouping obtained in this manner was used to perform the enrichment against KEs to avoid multiple testing against the same gene set as well as to enhance the network presentation of the pulmonary fibrosis (PF) AOP network. The unweighted PF AOP network was generated using gephi (107) by importing a graphml file generated with the function *graph_from_edgelist* from R package igraph (108). KE groups from the clustering were added as attributes to the nodes and used to merge redundant nodes in gephi. Similarly, AOPs each KE is associated to were added as attributes and used to colour the nodes.

Characterisation of the AOP-fingerprints

Transcriptomics data: *In vivo* and *in vitro* transcriptomics data from MWCNT (Mitsui-7) exposures were selected from a previously published collection by Saarimäki et al. (67). The original data sets are available under GEO accession number GSE29042 (*in vivo*) and ArrayExpress entry EMTAB6396 (*in vitro*), while the preprocessed data is available on Zenodo (<https://doi.org/10.5281/zenodo.6425445>). The *in vivo* data set comprises multiple doses and time points, while the *in vitro* data contains a single dose and time point exposure on four distinct cell lines representative different cell types of the lung. In each case, differentially expressed genes (filtered by an absolute fold change > 1.5 and Benjamini & Hochberg adjusted p-value < 0.05) obtained from Zenodo (<https://doi.org/10.5281/zenodo.6425445>) for each distinct comparison (i.e., combination of each dose and time point vs. control *in vivo* and separate cell lines *in vitro*) were pooled together to obtain a distinct mechanism of action of the exposure *in vivo* and *in vitro*, respectively.

AOP fingerprints: To produce the AOP fingerprint for the MWCNT exposures, enrichment analysis was performed using the Fisher’s exact test as implemented in the function *enrich* from R package bc3net (106) separately against the AOP-associated gene lists and the KE-associated gene lists (KEs linked to the same set of genes were grouped to avoid multiple tests against the same set). An AOP was considered significantly enriched when the AOP itself and at least 33% (or minimum of 2 KEs when the length of the AOP was less than 6) of its KEs were enriched with an FDR-corrected p-value < 0.05 .

Selection and testing of AOP-driven biomarkers

Gene prioritisation: All human PPI edges were extracted from the UKS and used to create a robust gene – gene network. PPI/gene (product) interaction information can vary across data sources as well as the covered genes may differ. In addition, there is an innate bias in the data, where more data sources are available for “more investigated” genes and gene products. Because of this, we decided not to apply a global threshold on how many sources need to support an edge (109), but instead, we applied a local threshold. This ensures that also less investigated genes and gene products will be retained in the final robust gene – gene network, but their edges are less penalised by the number of supporting edges, than highly covered gene (product) nodes. For each node, the mean number of sources supporting all its connecting edges was estimated and only edges with at least the “mean number of sources” for a node were retained, which needed to be true for at least one of the nodes making up an edge. This was only done for GENE nodes, which were flagged as protein coding in Ensembl. The final robust human gene – gene network, consisted of 20 260 nodes, 806 250 edges, a network density of 0.0039. Due to the significant lower number of available

sources for transcription factor – gene (product) data, all available sources were kept and scored equally. These edges were used to create a directed gene – gene network, consisting of 18 754 nodes, 363 649 edges and a network density of 0.001. On the so created gene – gene networks, for each node its degree, betweenness, eigenvector, and closeness centrality were estimated with NetworkX (110). These measures were then used to rank the genes linked to the KEs in the context of individual KEs. The gene list for each KE was ranked based on each of the centrality measures (degree, betweenness, closeness and eigenvector centrality) individually from most central to the least. The ranked lists were combined using the borda method as implemented in the function *Borda* in R package TopKLists (111).

Biomarker selection: The gene centrality-based ranking was then supplemented by a specificity ranking for the KEs of AOPs related to pulmonary fibrosis (PF). We calculated a specificity score for the genes in the context of the KEs by dividing the occurrence of the gene in the KEs of PF AOPs by their occurrence in the KEs of other AOPs. Similar score was calculated at the level of AOPs (occurrence in PF AOPs/occurrence other AOPs), as we were looking for universal PF biomarkers (i.e., prioritising those that would be present in as many of the six PF AOPs as possible) while also being as specific as possible to PF. These ranks were again combined by the function *Borda* from R package TopKLists (111), and a final round of the borda method was applied to combine the lists of genes from each KE into one PF rank. We complemented the final rank with experimental evidence. We assessed whether the genes were differentially expressed in the Mitsui-7 exposures *in vivo* and *in vitro*. We also evaluated whether they were dose-dependently altered in the *in vivo* data as well as in an additional *in vitro* data set on Mitsui-7 exposure of PMA-differentiated macrophages (originally published in Saarimäki *et al.* (58) and the preprocessed data available as GSE146708 in the previously published collection (67) available in <https://doi.org/10.5281/zenodo.6425445>). The dose-response modelling of the *in vivo* (GSE29042) and *in vitro* (GSE146708) datasets was performed by following the strategy implemented in the BMDx tool (112). Particularly, for each gene present in the dataset, multiple models were fitted including linear, second order polynomial, hill, power, and exponential model. For each gene, the optimal model was selected as the one with the lowest Akaike Information Criterion (AIC). Genes with an optimal model with lack-of-fit p-values lower than 0.1 were removed from the analysis. The effective doses (BMD, BMDL, and BMDU) were estimated under the assumption of constant variance and by using a BMRF factor of 1.349 (corresponding to a minimum of 10% of difference with respect to the controls). Genes were further filtered based on the predicted doses. Genes with BMD or BMDU values extrapolated higher than the highest exposure dose were filtered. Moreover, genes whose ratio between the predicted doses is higher than the suggested values (BMD/BMDL > 20, BMDU/BMD > 20, and BMDU/BMDL > 40) were removed from the analysis. Genes passing the filters were considered to be dose-dependently altered. At this stage, we also considered the measurability and feasibility of the gene in the selected macrophage model. For instance, numerous collagen-encoding genes were ranked high, but would not be a meaningful target in a macrophage model. Moreover, we aimed for a high coverage of PF KEs and aimed to select genes with high specificity scores. With these considerations, we selected a subset of the genes with the following priority: 1) genes that are deregulated both *in vivo* and *in vitro*, with most emphasis on dose-dependency, 2) genes that are deregulated *in vitro*, with most emphasis on dose-dependency and 3) genes that are not significantly differentially expressed but are dose-dependent. Finally, after this initial selection driven by the rank and experimental evidence, we included additional candidate biomarkers that had a lower rank

but were specific to KEs that would otherwise not have been covered by the selected candidates.

Cell culture: THP-1 cells (DSMZ no.: ACC 16) were grown in RPMI 1640 (Gibco, #21875) + 10% inactivated FBS (Gibco, #10270). Cells were cultivated in 75cm² culture flasks at 37°C with a humidified atmosphere of 5% CO₂. For all experiments, cells were seeded at a density of 1x10⁵ cells/ml in 96 well plates and differentiated for 48 hours with 25 nm PMA (phorbol-12-myristate-13-acetate, Sigma-Aldrich, #P1585). Cells were then left to rest for 24 hours in fresh media containing no PMA prior to bleomycin exposures.

Cell Viability Assay: THP-1 cells were exposed to 0-10 µg/ml of bleomycin ready-made solution (Sigma-Aldrich, #B7216) and 0-100 mg/ml of Triclosan (Sigma-Aldrich, #72779), for 6, 24 and 72 hours. A WST-1 assay was then used to measure cell viability. Briefly, 10 µl of cell proliferation reagent WST-1 (Roche, #11644807001) was added to each well. Cells were left to incubate with WST-1 for 3 hours in a 37°C, 5% CO₂ incubator. Absorbance at 450 nm was then measured with a Spark microplate reader (Tecan). Results of the cell viability assay are available in Supplementary file 2 and Figure S3.

RT-qPCR: For each time point of 6, 24 and 72 hours, THP-1 cells were exposed to 0, 2.5, 5, 10 and 100µg/ml of bleomycin ready-made solution (Sigma-Aldrich, #B7216). Media was removed and cells were washed briefly with 50 µl of PBS. 100 µl of lysis buffer from the QIAGEN RNeasy mini kit (Qiagen, #74104) was added to each well to lyse the cells. 3 wells (300 µl) were pooled to create one sample, and there were 5 samples for each concentration at each time point. Total RNA was then extracted from these samples using the QIAGEN RNeasy mini kit (Qiagen, #74104). cDNA was synthesised from 100 ng of RNA, using the High-capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, #4368813), according to manufacturer's instructions. Expression levels of target genes were determined by qRT-PCR using CFX96 Touch Real-Time PCR Detection System (BioRad) with 10 µl of iQ Multiplex Powermix (Bio-Rad, #1725849), 5 µl of cDNA diluted 5-fold, 2.5 µl of Nuclease-free (NF) water (not DEPC-Treated, ThermoFisher, #AM9930) in a 20 µl reaction, together with 2.5 µl of single (1 µl assay + 1.5 µl NF water) or multiplexed (0.5 µl of each assay) PrimePCR Probe Assays (Bio-Rad) as follows with single or multiplex reactions grouped in parentheses and formatted as *Gene/UniqueAssayID*: (*ACTB*/qHsaCEP0036280), (*SMAD7*/qHsaCEP0050142, *MMP9*/qHsaCIP0028098, *GDF15*/qHsaCEP0051579, *CTSK*/qHsaCIP0030907, *PLOD2*/qHsaCEP0052848), (*CXCL2*/qHsaCEP0058163, *LTBP4*/qHsaCEP0024931, *TGFB3*/qHsaCEP0058244, *RCN3*/qHsaCEP0057804, *MMP7*/qHsaCEP0052037), (*SPP1*/qHsaCEP0058179, *FN1*/qHsaCEP0050873, *LTBP3*/qHsaCEP0053782, *RSAD2*/qHsaCIP0031596, *CCL7*/qHsaCEP0058033), (*IL8*/qHsaCEP0053894, *MMP19*/qHsaCEP0051244, *TWIST1*/qHsaCEP0051221, *PLK3*/qHsaCIP0027687, *CXCL10*/qHsaCEP0053880), (*LOX*/qHsaCEP0050731, *PTX3*/qHsaCEP0033071, *TGFB1*/qHsaCIP0030973), *CCL2*/qHsaCIP0028103, *TGFB1*/qHsaCIP0030973).

Fold change (FC) values from RT-qPCR data were calculated using the comparative CT(2^{-(ddCt)}) method (113). The FC values were log₂ transformed (log₂(FC)). For each gene and for each concentration, an outlier detection was performed by removing all the samples with log₂(FC) values above or below the 75th and 25th percentiles of the distribution. Ct values, dCt values, FC values and log₂(FC) values are available in Supplementary file 2 along with ANOVA tables and tukey HSD posthoc test results. The log₂FC expression of the genes as compared to the untreated controls are plotted in Figure S4.

Dose-dependent modelling: A dose-response analysis of the log₂(FC) values derived from the PCR experiments was performed. For each gene, multiple models were fitted, including linear, hill, power, polynomial, exponential, log-logistic, Weibull, and Michaelis-Mentel models. The optimal model was selected as the one with the lowest AIC. The BMD estimation was performed under the assumption of constant variance. The BMR was identified by means of the standard deviation approach with a BMRF of 1.349. Only genes with lack-of-fit p-value >0.10 and with estimated BMD, BMDL and BMDU values were considered relevant.

Data Availability

The full list of KE annotation is available from the corresponding author upon request.

Acknowledgements

This work received funding from the EU Horizon 2020 project NanoSolveIT (grant agreement no. 814572) as well as the Academy of Finland (grant agreement no. 322761). Saarimäki was supported by the Emil Aaltonen Foundation.

References

1. Waters MD, Fostel JM. Toxicogenomics and systems toxicology: aims and prospects. *Nat Rev Genet.* 2004 Dec;5(12):936–48.
2. Liu Z, Huang R, Roberts R, Tong W. Toxicogenomics: A 2020 vision. *Trends Pharmacol Sci.* 2019 Feb;40(2):92–103.
3. Marx-Stoelting P, Braeuning A, Buhrke T, Lampen A, Niemann L, Oelgeschlaeger M, et al. Application of omics data in regulatory toxicology: report of an international BfR expert workshop. *Arch Toxicol.* 2015 Nov;89(11):2177–84.
4. Paananen J, Fortino V. An omics perspective on drug target discovery platforms. *Brief Bioinformatics.* 2020 Dec 1;21(6):1937–53.
5. Mech A, Gottardo S, Amenta V, Amodio A, Belz S, Bøwadt S, et al. Safe- and sustainable-by-design: The case of Smart Nanomaterials. A perspective based on a European workshop. *Regul Toxicol Pharmacol.* 2022 Feb;128:105093.
6. Richard AM, Huang R, Waidyanatha S, Shinn P, Collins BJ, Thillainadarajah I, et al. The tox21 10K compound library: collaborative chemistry advancing toxicology. *Chem Res Toxicol.* 2021 Feb 15;34(2):189–216.
7. Krewski D, Acosta D, Andersen M, Anderson H, Bailar JC, Boekelheide K, et al. Toxicity testing in the 21st century: a vision and a strategy. *J Toxicol Environ Health B Crit Rev.* 2010 Feb;13(2–4):51–138.
8. OECD. Guidance document for the use of adverse outcome pathways in developing integrated approaches to testing and assessment (IATA). OECD; 2017.

9. Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, et al. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem*. 2010 Mar;29(3):730–41.
10. Clerbaux L-A, Amigó N, Amorim MJ, Bal-Price A, Batista Leite S, Beronius A, et al. COVID-19 through Adverse Outcome Pathways: Building networks to better understand the disease - 3rd CIAO AOP Design Workshop. *ALTEX*. 2022 Jan 3;39(2):322–335.
11. Nymark P, Sachana M, Leite SB, Sund J, Krebs CE, Sullivan K, et al. Systematic Organization of COVID-19 Data Supported by the Adverse Outcome Pathway Framework. *Front Public Health*. 2021 May 19;9:638605.
12. Tollefsen KE, Scholz S, Cronin MT, Edwards SW, de Knecht J, Crofton K, et al. Applying Adverse Outcome Pathways (AOPs) to support Integrated Approaches to Testing and Assessment (IATA). *Regul Toxicol Pharmacol*. 2014 Dec;70(3):629–40.
13. Halappanavar S, van den Brule S, Nymark P, Gaté L, Seidel C, Valentino S, et al. Adverse outcome pathways as a tool for the design of testing strategies to support the safety assessment of emerging advanced materials at the nanoscale. Part Fibre Toxicol. 2020 May 25;17(1):16.
14. Fortino V, Kinaret PAS, Fratello M, Serra A, Saarimäki LA, Gallud A, et al. Biomarkers of nanomaterials hazard from multi-layer data. *Nat Commun*. 2022 Dec;13(1):3798.
15. Montaner J, Ramiro L, Simats A, Tiedt S, Makris K, Jickling GC, et al. Multilevel omics for the discovery of biomarkers and therapeutic targets for stroke. *Nat Rev Neurol*. 2020 May;16(5):247–64.
16. Federico A, Serra A, Ha MK, Kohonen P, Choi J-S, Liampa I, et al. Transcriptomics in toxicogenomics, part II: preprocessing and differential expression analysis for high quality data. *Nanomaterials (Basel)*. 2020 May 8;10(5).
17. Kinaret PAS, Serra A, Federico A, Kohonen P, Nymark P, Liampa I, et al. Transcriptomics in toxicogenomics, part I: experimental design, technologies, publicly available data, and regulatory aspects. *Nanomaterials (Basel)*. 2020 Apr 15;10(4).
18. Serra A, Fratello M, Cattelani L, Liampa I, Melagraki G, Kohonen P, et al. Transcriptomics in toxicogenomics, part III: data modelling for risk assessment. *Nanomaterials (Basel)*. 2020 Apr 8;10(4).
19. Brockmeier EK, Hodges G, Hutchinson TH, Butler E, Hecker M, Tollefsen KE, et al. The role of omics in the application of adverse outcome pathways for chemical risk assessment. *Toxicol Sci*. 2017 Aug 1;158(2):252–62.
20. Perkins EJ, Woolard EA, Garcia-Reyero N. Integration of adverse outcome pathways, causal networks and 'omics to support chemical hazard assessment. *Front Toxicol*. 2022 Mar 24;4:786057.

21. Labib S, Williams A, Yauk CL, Nikota JK, Wallin H, Vogel U, et al. Nano-risk Science: application of toxicogenomics in an adverse outcome pathway framework for risk assessment of multi-walled carbon nanotubes. *Part Fibre Toxicol*. 2016 Mar 15;13:15.
22. Jin Y, Qi G, Shou Y, Li D, Liu Y, Guan H, et al. High throughput data-based, toxicity pathway-oriented development of a quantitative adverse outcome pathway network linking AHR activation to lung damages. *J Hazard Mater*. 2022 Mar 5;425:128041.
23. Guan R, Li N, Wang W, Liu W, Li X, Zhao C. The adverse outcome pathway (AOP) of estrogen interference effect induced by triphenyl phosphate (TPP): Integrated multi-omics and molecular dynamics approaches. *Ecotoxicol Environ Saf*. 2022 Mar 7;234:113387.
24. Saarimäki LA, Melagraki G, Afantitis A, Lynch I, Greco D. Prospects and challenges for FAIR toxicogenomics data. *Nat Nanotechnol*. 2022 Jan;17(1):17–8.
25. Martens M, Verbruggen T, Nymark P, Grafström R, Burgoon LD, Aladjov H, et al. Introducing WikiPathways as a Data-Source to Support Adverse Outcome Pathways for Regulatory Risk Assessment of Chemicals and Nanomaterials. *Front Genet*. 2018 Dec 21;9:661.
26. Martens M, Evelo CT, Willighagen EL. Providing Adverse Outcome Pathways from the AOP-Wiki in a Semantic Web Format to Increase Usability and Accessibility of the Content. *Appl In Vitro Toxicol*. 2022 Mar 1;8(1):2–13.
27. Nymark P, Rieswijk L, Ehrhart F, Jeliaskova N, Tsiliki G, Sarimveis H, et al. A data fusion pipeline for generating and enriching adverse outcome pathway descriptions. *Toxicol Sci*. 2018 Mar 1;162(1):264–75.
28. Jagiello K, Halappanavar S, Rybińska-Fryca A, Williams A, Vogel U, Puzyn T. Transcriptomics-Based and AOP-Informed Structure-Activity Relationships to Predict Pulmonary Pathology Induced by Multiwalled Carbon Nanotubes. *Small*. 2021 Apr;17(15):e2003465.
29. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, Wiegiers J, Wiegiers TC, et al. Comparative Toxicogenomics Database (CTD): update 2021. *Nucleic Acids Res*. 2021 Jan 8;49(D1):D1138–43.
30. Tong M, Neusner A, Longato L, Lawton M, Wands JR, de la Monte SM. Nitrosamine exposure causes insulin resistance diseases: relevance to type 2 diabetes mellitus, non-alcoholic steatohepatitis, and Alzheimer’s disease. *J Alzheimers Dis*. 2009;17(4):827–44.
31. Ceballos DM, Fellows KM, Evans AE, Janulewicz PA, Lee EG, Whittaker SG. Perchloroethylene and dry cleaning: it’s time to move the industry to safer alternatives. *Front Public Health*. 2021 Mar 5;9:638082.

32. Cave M, Falkner KC, McClain C. Occupational and environmental hepatotoxicity. *Zakim and boyer's hepatology*. Elsevier; 2012. p. 476–92.
33. Chow W-H, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. *Nat Rev Urol*. 2010 May;7(5):245–57.
34. Krenkel O, Mossanen JC, Tacke F. Immune mechanisms in acetaminophen-induced acute liver failure. *Hepatobiliary Surg Nutr*. 2014 Dec;3(6):331–43.
35. Jia M, Wang Y, Wang D, Teng M, Yan J, Yan S, et al. The effects of hexaconazole and epoxiconazole enantiomers on metabolic profile following exposure to zebrafish (*Danio rerio*) as well as the histopathological changes. *Chemosphere*. 2019 Jul;226:520–33.
36. Lee W-S, Kim J. Statin-induced liver and muscle toxicities. *Mol Cell Toxicol*. 2019 Jan;15(1):9–17.
37. Warner GR, Mourikes VE, Neff AM, Brehm E, Flaws JA. Mechanisms of action of agrochemicals acting as endocrine disrupting chemicals. *Mol Cell Endocrinol*. 2020 Feb 15;502:110680.
38. Brander SM, He G, Smalling KL, Denison MS, Cherr GN. The in vivo estrogenic and in vitro anti-estrogenic activity of permethrin and bifenthrin. *Environ Toxicol Chem*. 2012 Dec;31(12):2848–55.
39. Moore PD, Yedjou CG, Tchounwou PB. Malathion-induced oxidative stress, cytotoxicity, and genotoxicity in human liver carcinoma (HepG2) cells. *Environ Toxicol*. 2010 Jun;25(3):221–6.
40. Xiong J, Tian L, Qiu Y, Sun D, Zhang H, Wu M, et al. Evaluation on the thyroid disrupting mechanism of malathion in Fischer rat thyroid follicular cell line FRTL-5. *Drug Chem Toxicol*. 2018 Oct;41(4):501–8.
41. Lal B, Sarang MK, Kumar P. Malathion exposure induces the endocrine disruption and growth retardation in the catfish, *Clarias batrachus* (Linn.). *Gen Comp Endocrinol*. 2013 Jan 15;181:139–45.
42. Grasselli F, Bussolati S, Ramoni R, Grolli S, Basini G. Simazine, a triazine herbicide, disrupts swine granulosa cell functions. *Anim Reprod*. 2018 Aug 16;15(1):3–11.
43. Tu W, Xu C, Jin Y, Lu B, Lin C, Wu Y, et al. Permethrin is a potential thyroid-disrupting chemical: In vivo and in silico evidence. *Aquat Toxicol*. 2016 Jun;175:39–46.
44. Kim S-S, Lee R-D, Lim K-J, Kwack S-J, Rhee G-S, Seok J-H, et al. Potential estrogenic and antiandrogenic effects of permethrin in rats. *J Reprod Dev*. 2005 Apr;51(2):201–10.

45. Yang JS, Qi W, Farias-Pereira R, Choi S, Clark JM, Kim D, et al. Permethrin and ivermectin modulate lipid metabolism in steatosis-induced HepG2 hepatocyte. *Food Chem Toxicol*. 2019 Mar;125:595–604.
46. Ravichandran J, Karthikeyan BS, Samal A. Investigation of a derived adverse outcome pathway (AOP) network for endocrine-mediated perturbations. *Sci Total Environ*. 2022 Jun 20;826:154112.
47. Spînu N, Cronin MTD, Lao J, Bal-Price A, Campia I, Enoch SJ, et al. Probabilistic modelling of developmental neurotoxicity based on a simplified adverse outcome pathway network. *Comput Toxicol*. 2022 Feb;21:100206.
48. Zhang T, Wang S, Li L, Zhu A, Wang Q. Associating diethylhexyl phthalate to gestational diabetes mellitus via adverse outcome pathways using a network-based approach. *Sci Total Environ*. 2022 Jun 10;824:153932.
49. Arnesdotter E, Spinu N, Firman J, Ebbrell D, Cronin MTD, Vanhaecke T, et al. Derivation, characterisation and analysis of an adverse outcome pathway network for human hepatotoxicity. *Toxicology*. 2021 Jul 10;459:152856.
50. Knapen D, Angrish MM, Fortin MC, Katsiadaki I, Leonard M, Margiotta-Casaluci L, et al. Adverse outcome pathway networks I: Development and applications. *Environ Toxicol Chem*. 2018 Jun;37(6):1723–33.
51. Villeneuve DL, Angrish MM, Fortin MC, Katsiadaki I, Leonard M, Margiotta-Casaluci L, et al. Adverse outcome pathway networks II: Network analytics. *Environ Toxicol Chem*. 2018 Jun;37(6):1734–48.
52. Thannickal VJ, Toews GB, White ES, Lynch JP, Martinez FJ. Mechanisms of pulmonary fibrosis. *Annu Rev Med*. 2004;55:395–417.
53. Bonner JC. Lung fibrotic responses to particle exposure. *Toxicol Pathol*. 2007 Jan;35(1):148–53.
54. Wynn TA. Cellular and molecular mechanisms of fibrosis. *J Pathol*. 2008 Jan;214(2):199–210.
55. McGroder CF, Zhang D, Choudhury MA, Salvatore MM, D’Souza BM, Hoffman EA, et al. Pulmonary fibrosis 4 months after COVID-19 is associated with severity of illness and blood leucocyte telomere length. *Thorax*. 2021 Dec;76(12):1242–5.
56. Kinaret PAS, Del Giudice G, Greco D. Covid-19 acute responses and possible long term consequences: What nanotoxicology can teach us. *Nano Today*. 2020 Dec;35:100945.
57. Hama Amin BJ, Kakamad FH, Ahmed GS, Ahmed SF, Abdulla BA, Mohammed SH, et al. Post COVID-19 pulmonary fibrosis; a meta-analysis study. *Ann Med Surg (Lond)*. 2022 May;77:103590.

58. Saarimäki LA, Kinaret PAS, Scala G, del Giudice G, Federico A, Serra A, et al. Toxicogenomics analysis of dynamic dose-response in macrophages highlights molecular alterations relevant for multi-walled carbon nanotube-induced lung fibrosis. *NanoImpact*. 2020 Nov;100274.
59. Scala G, Kinaret P, Marwah V, Sund J, Fortino V, Greco D. Multi-omics analysis of ten carbon nanomaterials effects highlights cell type specific patterns of molecular regulation and adaptation. *NanoImpact*. 2018 Jul;11:99–108.
60. Kinaret P, Marwah V, Fortino V, Ilves M, Wolff H, Ruokolainen L, et al. Network analysis reveals similar transcriptomic responses to intrinsic properties of carbon nanomaterials in vitro and in vivo. *ACS Nano*. 2017 Apr 25;11(4):3786–96.
61. Rydman EM, Ilves M, Vanhala E, Vippola M, Lehto M, Kinaret PAS, et al. A Single Aspiration of Rod-like Carbon Nanotubes Induces Asbestos-like Pulmonary Inflammation Mediated in Part by the IL-1 Receptor. *Toxicol Sci*. 2015 Sep;147(1):140–55.
62. Kinaret P, Ilves M, Fortino V, Rydman E, Karisola P, Lähde A, et al. Inhalation and Oropharyngeal Aspiration Exposure to Rod-Like Carbon Nanotubes Induce Similar Airway Inflammation and Biological Responses in Mouse Lungs. *ACS Nano*. 2017 Jan 24;11(1):291–303.
63. Søs Poulsen S, Jacobsen NR, Labib S, Wu D, Husain M, Williams A, et al. Transcriptomic analysis reveals novel mechanistic insight into murine biological responses to multi-walled carbon nanotubes in lungs and cultured lung epithelial cells. *PLoS ONE*. 2013 Nov 19;8(11):e80452.
64. Rahman L, Jacobsen NR, Aziz SA, Wu D, Williams A, Yauk CL, et al. Multi-walled carbon nanotube-induced genotoxic, inflammatory and pro-fibrotic responses in mice: Investigating the mechanisms of pulmonary carcinogenesis. *Mutat Res Genet Toxicol Environ Mutagen*. 2017 Nov;823:28–44.
65. Guo NL, Wan Y-W, Denvir J, Porter DW, Pacurari M, Wolfarth MG, et al. Multiwalled carbon nanotube-induced gene signatures in the mouse lung: potential predictive value for human lung cancer risk and prognosis. *J Toxicol Environ Health Part A*. 2012;75(18):1129–53.
66. Scala G, Marwah V, Kinaret P, Sund J, Fortino V, Greco D. Integration of genome-wide mRNA and miRNA expression, and DNA methylation data of three cell lines exposed to ten carbon nanomaterials. *Data Brief*. 2018 Aug;19:1046–57.
67. Saarimäki LA, Federico A, Lynch I, Papadiamantis AG, Tsoumanis A, Melagraki G, et al. Manually curated transcriptomics data collection for toxicogenomic assessment of engineered nanomaterials. *Sci Data*. 2021 Feb 8;8(1):49.
68. Barbarino M, Giordano A. Assessment of the carcinogenicity of carbon nanotubes in the respiratory system. *Cancers (Basel)*. 2021 Mar 15;13(6).

69. Rittinghausen S, Hackbarth A, Creutzenberg O, Ernst H, Heinrich U, Leonhardt A, et al. The carcinogenic effect of various multi-walled carbon nanotubes (MWCNTs) after intraperitoneal injection in rats. *Part Fibre Toxicol*. 2014 Nov 20;11:59.
70. Gupta SS, Singh KP, Gupta S, Dusinska M, Rahman Q. Do carbon nanotubes and asbestos fibers exhibit common toxicity mechanisms? *Nanomaterials (Basel)*. 2022 May 17;12(10).
71. Boyles MSP, Young L, Brown DM, MacCalman L, Cowie H, Moisola A, et al. Multi-walled carbon nanotube induced frustrated phagocytosis, cytotoxicity and pro-inflammatory conditions in macrophages are length dependent and greater than that of asbestos. *Toxicol In Vitro*. 2015 Oct;29(7):1513–28.
72. Porter DW, Hubbs AF, Mercer RR, Wu N, Wolfarth MG, Sriram K, et al. Mouse pulmonary dose- and time course-responses induced by exposure to multi-walled carbon nanotubes. *Toxicology*. 2010 Mar 10;269(2–3):136–47.
73. Izbicki G, Segel MJ, Christensen TG, Conner MW, Breuer R. Time course of bleomycin-induced lung fibrosis. *Int J Exp Pathol*. 2002 Jun;83(3):111–9.
74. Hill AB. THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? *Proc R Soc Med*. 1965 May;58:295–300.
75. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol*. 2015 Sep 30;12(1):14.
76. Yue X, Shan B, Lasky JA. TGF- β : Titan of Lung Fibrogenesis. *Curr Enzym Inhib*. 2010 Jul 1;6(2).
77. Annes JP, Munger JS, Rifkin DB. Making sense of latent TGFbeta activation. *J Cell Sci*. 2003 Jan 15;116(Pt 2):217–24.
78. Nummela P, Lammi J, Soikkeli J, Saksela O, Laakkonen P, Hölttä E. Transforming growth factor beta-induced (TGFBI) is an anti-adhesive protein regulating the invasive growth of melanoma cells. *Am J Pathol*. 2012 Apr;180(4):1663–74.
79. Yan X, Liu Z, Chen Y. Regulation of TGF-beta signaling by Smad7. *Acta Biochim Biophys Sin (Shanghai)*. 2009 Apr;41(4):263–72.
80. Bachy S, Wu Z, Gamradt P, Thierry K, Milani P, Chlasta J, et al. β ig-h3-structured collagen alters macrophage phenotype and function in pancreatic cancer. *iScience*. 2022 Feb 18;25(2):103758.
81. Hou J, Shi J, Chen L, Lv Z, Chen X, Cao H, et al. M2 macrophages promote myofibroblast differentiation of LR-MSCs and are associated with pulmonary fibrogenesis. *Cell Commun Signal*. 2018 Nov 23;16(1):89.

82. Wang H, Gao Y, Wang L, Yu Y, Zhang J, Liu C, et al. Lung specific homing of diphenylethylideneiodonium chloride improves pulmonary fibrosis by inhibiting macrophage M2 metabolic program. *J Advanc Res*. 2022 Apr;
83. Hammond ME, Lapointe GR, Feucht PH, Hilt S, Gallegos CA, Gordon CA, et al. IL-8 induces neutrophil chemotaxis predominantly via type I IL-8 receptors. *J Immunol*. 1995 Aug 1;155(3):1428–33.
84. De Filippo K, Dudeck A, Hasenberg M, Nye E, van Rooijen N, Hartmann K, et al. Mast cell and macrophage chemokines CXCL1/CXCL2 control the early stage of neutrophil recruitment during tissue inflammation. *Blood*. 2013 Jun 13;121(24):4930–7.
85. Cheng JW, Sadeghi Z, Levine AD, Penn MS, von Recum HA, Caplan AI, et al. The role of CXCL12 and CCL7 chemokines in immune regulation, embryonic development, and tissue regeneration. *Cytokine*. 2014 Oct;69(2):277–83.
86. Braga TT, Agudelo JSH, Camara NOS. Macrophages during the fibrotic process: M2 as friend and foe. *Front Immunol*. 2015 Nov 25;6:602.
87. Dong J, Ma Q. Macrophage polarization and activation at the interface of multi-walled carbon nanotube-induced pulmonary inflammation and fibrosis. *Nanotoxicology*. 2018 Jan 16;12(2):153–68.
88. Kinaret PAS, Scala G, Federico A, Sund J, Greco D. Carbon nanomaterials promote M1/M2 macrophage activation. *Small*. 2020 May;16(21):e1907609.
89. Meng J, Li X, Wang C, Guo H, Liu J, Xu H. Carbon nanotubes activate macrophages into a M1/M2 mixed status: recruiting naïve macrophages and supporting angiogenesis. *ACS Appl Mater Interfaces*. 2015 Feb 11;7(5):3180–8.
90. Löffek S, Schilling O, Franzke CW. Series “matrix metalloproteinases in lung health and disease”: Biological role of matrix metalloproteinases: a critical balance. *Eur Respir J*. 2011 Jul;38(1):191–208.
91. Pardo A, Cabrera S, Maldonado M, Selman M. Role of matrix metalloproteinases in the pathogenesis of idiopathic pulmonary fibrosis. *Respir Res*. 2016 Mar 4;17:23.
92. Craig VJ, Zhang L, Hagood JS, Owen CA. Matrix metalloproteinases as therapeutic targets for idiopathic pulmonary fibrosis. *Am J Respir Cell Mol Biol*. 2015 Nov;53(5):585–600.
93. Yu G, Kovkarova-Naumovski E, Jara P, Parwani A, Kass D, Ruiz V, et al. Matrix metalloproteinase-19 is a key regulator of lung fibrosis in mice and humans. *Am J Respir Crit Care Med*. 2012 Oct 15;186(8):752–62.
94. Federico A, Fratello M, Scala G, Möbus L, Pavel A, Del Giudice G, et al. Integrated Network Pharmacology Approach for Drug Combination Discovery: A Multi-Cancer Case Study. *Cancers (Basel)*. 2022 Apr 18;14(8).

95. Pavel A, del Giudice G, Federico A, Di Lieto A, Kinaret PAS, Serra A, et al. Integrated network analysis reveals new genes suggesting COVID-19 chronic effects and treatment. *Brief Bioinformatics*. 2021;
96. Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, et al. WikiPathways: connecting communities. *Nucleic Acids Res*. 2021 Jan 8;49(D1):D613–21.
97. Köhler S, Gargano M, Matentzoglou N, Carmody LC, Lewis-Smith D, Vasilevsky NA, et al. The human phenotype ontology in 2021. *Nucleic Acids Res*. 2021 Jan 8;49(D1):D1207–17.
98. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000 Jan 1;28(1):27–30.
99. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The reactome pathway knowledgebase 2022. *Nucleic Acids Res*. 2022 Jan 7;50(D1):D687–92.
100. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*. 2000 May;25(1):25–9.
101. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022. *Nucleic Acids Res*. 2022 Jan 7;50(D1):D988–95.
102. Sund J, Deceuninck P. EURL ECVAM library of reference chemicals. European Commission, Joint Research Centre (JRC) [Internet]. 2021 [cited 2022 Jul 6]. Available from: <http://data.europa.eu/89h/92614229-d020-4d96-941c-c9604e525c9e>
103. Wegner S, Browne P, Dix D. Identifying reference chemicals for thyroid bioactivity screening. *Reprod Toxicol*. 2016 Oct;65:402–13.
104. Kleinstreuer NC, Ceger P, Watt ED, Martin M, Houck K, Browne P, et al. Development and validation of a computational model for androgen receptor activity. *Chem Res Toxicol*. 2017 Apr 17;30(4):946–64.
105. Browne P, Judson RS, Casey WM, Kleinstreuer NC, Thomas RS. Screening chemicals for estrogen receptor bioactivity using a computational model. *Environ Sci Technol*. 2015 Jul 21;49(14):8804–14.
106. de Matos Simoes R, Emmert-Streib F. Bagging statistical network inference from large-scale gene expression data. *PLoS ONE*. 2012 Mar 30;7(3):e33624.
107. Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media*. 2009 Mar 19;

108. Csárdi G, Nepusz T. The igraph software package for complex network research. undefined. 2006;
109. Pavel A, Serra A, Cattelani L, Federico A, Greco D. Network analysis of microarray data. *Methods Mol Biol.* 2022;2401:161–86.
110. Hagberg AA, Schult DA, Swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX. Pasadena, CA, USA: SciPy 2008; 2008. p. 11–5.
111. Schimek MG, Budinská E, Kugler KG, Švendová V, Ding J, Lin S. TopKLists: a comprehensive R package for statistical inference, stochastic aggregation, and visualization of multiple omics ranked lists. *Stat Appl Genet Mol Biol.* 2015 Jun;14(3):311–6.
112. Serra A, Saarimäki LA, Fratello M, Marwah VS, Greco D. BMDx: a graphical Shiny application to perform Benchmark Dose analysis for transcriptomics data. *Bioinformatics.* 2020 May 1;36(9):2932–3.
113. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001 Dec;25(4):402–8.