Title

Advancing chemical hazard assessment with decision analysis: A case study on lithium-ion and redox flow batteries used for energy storage.

Permalink

https://escholarship.org/uc/item/6733q78z

Authors

He, Haoyang
Tian, Shan
Glaubensklee, Chris
et al.

Publication Date

2022-09-01

DOI

10.1016/j.jhazmat.2022.129301

Journal of Hazardous Materials 437 (2022) 129301

Contents lists available at ScienceDirect

Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

Research Paper

Advancing chemical hazard assessment with decision analysis: A case study on lithium-ion and redox flow batteries used for energy storage

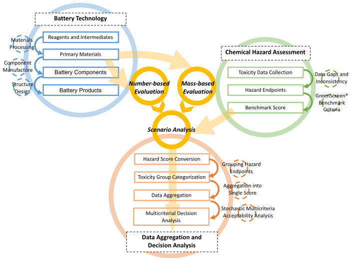
Haoyang He [a], Shan Tian [b,c], Chris Glaubensklee [a], Brian Tarroja [c,d], Scott Samuelsen [b,c], Oladele A. Ogunseitan [e], Julie M. Schoenung [a,*]

[a] *Department of Materials Science and Engineering, University of California, Irvine, CA, United States*
[b] *Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, United States*
[c] *Advanced Power and Energy Program, University of California, Irvine, CA, United States*
[d] *Department of Civil and Environmental Engineering, University of California, Irvine, CA, United States*
[e] *Department of Population Health and Disease Prevention, University of California, Irvine, CA, United States*

HIGHLIGHTS

- Chemical hazard assessment of substances used in six lithium ion and flow batteries.
- Materials of high concern and specific hazard endpoints identified for each battery.
- Data inconsistencies, data gaps and uncertainty evaluated for each material assessed.
- Data aggregation approach developed for both material and product based assessments.
- Multi-criteria decision analysis applied to identify safer battery alternatives.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Dr. Y. Deng

Keywords:
Battery energy storage
Chemical hazard assessment
Multi-criteria decision analysis
Toxicity
Human health impact

ABSTRACT

Batteries are important for promoting renewable energy, but, like most engineered products, they contain multiple hazardous materials. The purpose of this study is to evaluate industrial-scale batteries using GreenScreen® for Safer Chemicals, an established chemical hazard assessment (CHA) framework, and to develop a systematic, transparent methodology to quantify the CHA results, harmonize them, and aggregate them into single-value hazard scores, which can facilitate quantitative comparison and a robust evaluation of data gaps, inconsistencies, and uncertainty through the implementation of carefully selected scenarios and stochastic multicriteria acceptability analysis (SMAA). Using multiple authoritative toxicity data sources, six battery products are evaluated: three lithium-ion batteries (lithium iron phosphate, lithium nickel cobalt manganese hydroxide, and lithium manganese oxide), and three redox flow batteries (vanadium redox, zinc-bromine, and all-iron). The CHA results indicate that many materials in these batteries, including reagents and intermediates, inherently exhibit high hazard; therefore, safer materials should be identified and considered in future designs. The scenario analysis and SMAA, combined, provide a quantitative evaluation framework to support the decision-making needed to compare alternative technologies. Thus, this study highlights specific strategies to

---

* Corresponding author.
E-mail address: Julie.Schoenung@UCI.edu (J.M. Schoenung).

https://doi.org/10.1016/j.jhazmat.2022.129301
Received 7 April 2022; Received in revised form 18 May 2022; Accepted 2 June 2022

reduce the use of hazardous materials in complex engineered products before they are widely used in this rapidly-expanding industry sector.

1. Introduction

Concerns about environmental threats such as climate change, air pollution, and water pollution have motivated the deployment of alternative energy resources and the technologies that support them (Larcher and Tarascon, 2014). For example, California is searching for solutions through implementing policies such as the Renewables Portfolio Standard (RPS) and Senate Bill 100 which requires 60 % of retail electricity sales to be sourced from renewable resources by 2025 and 100 % of the electric demand to be met by zero-carbon electricity resources by 2045 (CA Senate Bill No.32, 2016; CA Senate Bill No.100, 2018). Meeting the RPS and SB 100 goals will likely depend strongly on the use of wind and solar energy (California Energy Commission, 2021). However, these resources can cause mismatching of electricity demand and supply due to the variability of their electricity generation profile. Energy storage can be an effective way to compensate for this mismatch and enable a high renewable penetration level on the electric grid, and in particular, battery storage technologies provide a high-efficiency solution towards this end. Currently, various battery technologies are being considered to fulfill such a role such as lithium-ion batteries and flow batteries. Of these two battery technologies, lithium-ion batteries are currently ubiquitous due to their high energy density enabling use in portable applications and generally lower cost annually per battery pack (Schmidt et al., 2019a; Wood III et al., 2015). In addition to enabling the proliferation of electric vehicles, lithium-ion batteries can be potential candidates for large scale grid-connected energy storage (Hiremath et al., 2015; Schmidt et al., 2019b; Tarroja et al., 2016). Although flow batteries are configured to have lower energy density compared to lithium-ion batteries, they have the potential to be better customized to provide services to a renewable grid due to their ability to independently configure their power and energy capacity. These benefits render flow batteries as a potential key component of renewable energy system infrastructure (Tian et al., 2021).

While renewable energy systems are considered to have lower environmental impacts compared to non-renewable alternatives, the manufacturing of batteries to scale-up their use is not burden-free. Sustainability assessment fulfills the key role of identifying the environmental trade-offs associated with the deployment of renewable energy systems. Research projects focusing on sustainability assessment of batteries are emerging under the pressure to promote cleaner energy use. For example, life cycle assessments have been conducted to investigate the environmental impact of implementing battery storage technologies (da Silva Lima et al., 2021; Ellingsen et al., 2014; Liang et al., 2017; Weber et al., 2018). Materials flow analysis has been applied to explore the resource criticality in the materials supply chain (Olivetti et al., 2017; Song et al., 2019). Risk assessment has been conducted to study the physical hazards (e.g., flammability) caused by chemical exposure to highlight the safety issues of battery use (Blum and Long, 2016; Chen et al., 2021; Kang et al., 2013; Kalhoff et al., 2015; Liu et al., 2018; Lisbona and Snee, 2011). In addition, strategies on sustainable waste management of end-of-life battery have also been investigated to facilitate the circular economy (Awasthi et al., 2021; Harper et al., 2019; Tenaw et al., 2021). However, most of the published research focused on relatively mature, commercially available technologies such as lead-acid batteries and conventional lithium-ion batteries. Additionally, previous studies have not systematically evaluated the inherent toxicity and potential hazard to human health and the environment associated with materials use throughout the battery production chain. Therefore, there is a need to conduct systematic hazard assessments for emerging energy storage technologies such as advanced lithium-ion and novel flow batteries to better support decision-making regarding the development and deployment of grid-connected energy storage.

In response to the current demand for safer and cleaner products before they are commercialized, chemical hazard assessment (CHA), which is designed to assess the toxicity and hazard of a chemical for a given application, is endorsed by legislative initiatives, governmental and non-governmental organizations, and research institutes (Department of Toxic Substances Control, 2021; National Research Council, 2014; United States Environmental Protection Agency, 2021a). The principle of CHA makes it a powerful tool to evaluate and identify chemicals of high concern used for the production of engineered products (Aschberger et al., 2017; Eisenberg et al., 2013; He et al., 2019a; Prat et al., 2016). Nevertheless, the current CHA framework is only designed to assess and compare individual chemicals, whereas engineered products such as batteries consist of complex materials, functional components, and of course the whole product. Aggregating hazard, toxicity and CHA-derived data for decision making at the component and/or product level is non-trivial with existing CHA tools and frameworks and is further compounded by data gaps and uncertainty. Methodological development for a CHA framework that can accommodate an expanded assessment boundary is urgently needed to facilitate product comparisons that account not only for conventional attributes such as performance and cost, but also potential impact to human health and the environment. In addition, strategies are need for addressing the data gaps, data inconsistencies and quantitative characterization of assumptions in such a framework. These uncertainties can be evaluated using scenarios and decision analysis methods.

Decision analysis is a viable approach to evaluate trade-offs when confronted with divergent results given the variety of hazard endpoints evaluated (Jacobs et al., 2016; OECD, 2021). A review of the literature shows few attempts to apply decision analysis in the CHA framework, using various methods and tools ranging from informal rules based on expert judgment to complex statistical methods (Linkov et al., 2009; Malloy et al., 2017; OECD, 2021). Multi-criteria decision analysis (MCDA) is a quantitative and systematic approach that assesses multiple criteria in decision making. While commonly used for business cases, MCDA has been only occasionally applied in sustainability assessments such as life cycle assessment, risk assessment, and alternative analysis, facilitating data aggregation, uncertainty analysis and comparison of alternatives (Cinelli et al., 2014; He et al., 2019b; Khakzad and Reniers, 2015; Khan et al., 2018; Zanghelini et al., 2018; Zheng et al., 2019). When applying MCDA in CHA, the model construction and problem scoping, the selection of proper MCDA methods, the rules for criteria evaluation, and the choices for weighting schemes, should be carefully and systematically considered.

To our knowledge, the current literature does not provide an up-to-date, systematic, and comprehensive hazard assessment of materials used in batteries that are essential to their functionality. Before batteries with various chemistries are installed extensively throughout the world, consideration of material hazard is urgently needed to minimize potential risk and exposure. Thus, the goal of this study is to investigate the inherent toxicity and hazard associated with the materials contained in battery products used for energy storage on the grid. Furthermore, a robust quantification and data aggregation approach coupled with MCDA is applied to integrate the results from various hazard endpoints and evaluate uncertainty through the application of scenario analysis. We systematically evaluated six types of battery technologies, which include three lithium-ion batteries (LIBs): lithium iron phosphate (LFP), lithium nickel cobalt manganese hydroxide (NCM), and lithium manganese oxide (LMO) (Majeau-Bettez et al., 2011; Notter et al., 2010); and three redox flow batteries (RFBs): vanadium redox (VRFB), zinc-bromine (ZBFB), and all-iron (IFB) (He et al., 2020). GreenScreen®

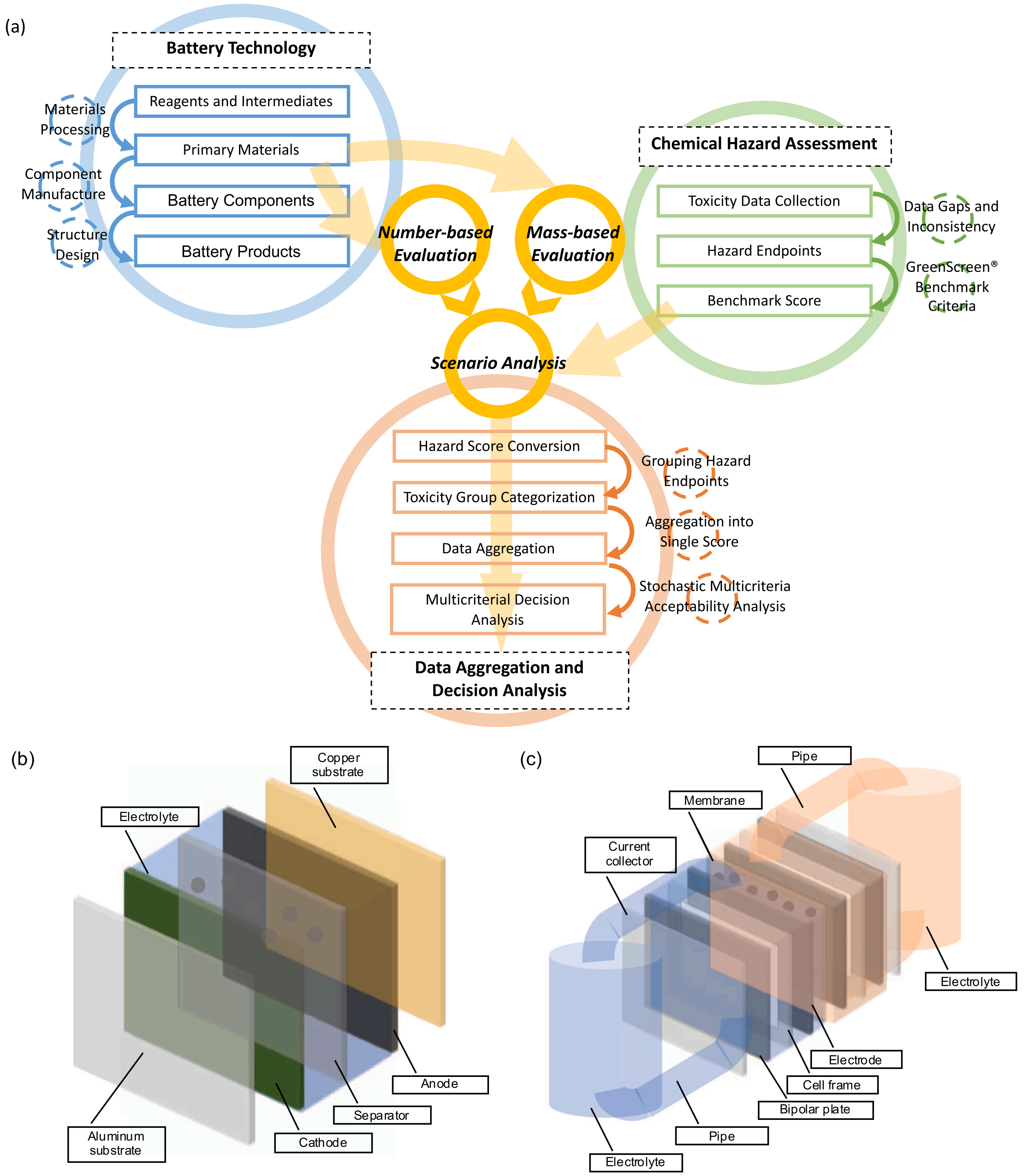


Fig. 1. (a) The integrated assessment framework used to evaluate toxicity and hazard for energy storage battery technologies. The structure and component illustrations for: (b) *lithium-ion* batteries and (c) *redox flow* batteries.

for Safer Chemicals version 1.4 (GreenScreen®) and stochastic multicriteria acceptability analysis (SMAA) are used for CHA and MCDA, respectively (Clean Production Action, 2021a; Lahdelma et al., 1998; Lahdelma and Salminen, 2001). The important contributions of our own study were two-fold: (1) a systematic evaluation of toxicity and hazard endpoints for primary materials and reagents/intermediates in these six batteries, and (2) development and implementation of a viable approach to quantify and aggregate the disparate toxicity and hazard endpoint data on materials into hazard scores for complex, multi-material, engineered products, allowing for the consideration of uncertainty through scenario analysis in combination with decision analysis methods to robustly yet transparently compare alternatives.

2. Materials and methods

As shown in Fig. 1**a**, the integrated assessment approach used in this study include: description of the components and materials from which the battery products are made; conducting the chemical hazard assessment (CHA); and developing a robust, yet systematic and transparent, assessment approach to aggregate the CHA data to the battery product level allowing for comparison using decision analysis methods. Specific strategies for developing the assessment framework include: harmonizing data inconsistencies among variable data sources and hazard endpoints, evaluating reagents and intermediates for polymeric materials and other compounds for which direct hazard data do not exist, converting qualitative hazard descriptions into quantitative hazard scores for numerical computation, comparing the material-specific

Table 1
The hazard endpoint information included in GreenScreen® for Safer Chemicals (Clean Production Action, 2021a).

Hazard Groups	Hazard Endpoints
Human Health Group I	Carcinogenicity (C), Mutagenicity & Genotoxicity (M), Reproductive Toxicity (R), Developmental Toxicity (D), Endocrine Activity (E)
Human Health Group II	Acute Mammalian Toxicity (AT), Systemic Toxicity & Organ Effects (ST-single), Neurotoxicity (N-single), Skin Irritation (IrS), Eye Irritation (IrE)
Human Health Group II*	Systemic Toxicity & Organ Effects* Repeated Exposure sub-endpoint (ST-repeated), Neurotoxicity * Repeated Exposure sub-endpoint (N-repeated), Skin Sensitization (SnS), Respiratory Sensitization (SnR)
Environmental Toxicity & Fate	Acute Aquatic Toxicity (AA), Chronic Aquatic Toxicity (CA), Persistence (P), Bioaccumulation (B)
Physical Hazards	Reactivity (Rx), Flammability (F)

results when aggregated either by hazard frequency or by material mass, and aggregating hazard scores for materials into hazard scores for battery products, so that the comprehensive hazard profile of each battery product can be compared by implementing multiple scenarios and multi-criteria decision analysis. Each step is described in detail below.

2.1. Materials in six selected energy storage battery products

In Fig. 1**b** and **c**, schematic diagrams illustrate the internal structures and major components for LIBs and RFBs, respectively. Specifications for the battery components and the associated mass fractions are provided in Tables S1 and S2 in the Appendix I-Supplemental Information (SI). To focus on components that are essential for battery function, peripheral components such as the battery management system, inverter, battery packaging, and accessories (shaded in gray in Tables S1 and S2) are not included in our assessments since these can vary significantly from manufacturer to manufacturer. Instead, our assessments only consider the functional components, i.e., the battery cell components, for all six battery types. The configurations for the three LIBs are similar to each other regardless of the different chemical types. The components of the cell stack for all three LIBs include electrodes categorized as cathodes and anodes, separators inserted to divide the two electrodes, and electrolytes between the cathode and anode (Zubi et al., 2018). Substrates used as current collectors are typically made of aluminum and copper. RFBs differ from LIBs in that the electrolytes are stored in external tanks that are connected to the cell stack via pipes and pumps, enabling physical separation of the power and energy capacity subsystems (Chalamala et al., 2014; Leung et al., 2012). Unlike LIBs in which the key materials that conduct the redox reactions are embedded in the cathode, the active species for RFBs are dissolved in the electrolytes (Noack et al., 2015). Other components inside the RFBs cell stack include the bipolar plate, cell frame, electrode, membrane, and current collector. In our case, the ZBFB does not include an additional electrode and membrane materials compared to VRFB and IFB due to the use of an integrated titanium bipolar plate instead of a more traditional carbon-based bipolar plate. For both LIBs and RFBs, the most robust industrial-scale life cycle inventory data were used as the basis for the evaluation in this study. Specifically, for LIBs, the inventory of primary materials was derived from published industrial-scale life cycle assessment (LCA) studies (Majeau-Bettez et al., 2011; Notter et al., 2010). These studies report the use of different cathode active materials, which distinguish the three LIBs. There are also different material choices for components such as the separator and electrode binder and solvent. Data on the primary materials in RFBs were collected from manufacturers, as detailed in the study by He et al. (2020). The main difference between the three RFBs is the different active materials used as electrolytes. Data from manufacturers indicate that material choices for the bipolar plate, cell frame, current collector, and tank also vary (He et al., 2020).

Table 2
Conversion of qualitative hazard endpoint descriptions to quantitative hazard scores.

Toxicity Group	Qualitative Hazard Endpoint Description	Hazard Score
Carcinogenic and Mutagenic Toxicity (C, M)	Strong evidence of negative studies, no hazard data, considered safe	1
	“Low (L)” hazard level specified by GreenScreen®	2
	“Moderate (M)” hazard level specified by GreenScreen®	3
	“High (H)” hazard level specified by GreenScreen®	4
Reproductive, Developmental, and Endocrine Toxicity (R, D, E)	Strong evidence of negative studies, no hazard data, considered safe	1
	“Low (L)” hazard level specified by GreenScreen®	2
	“Moderate (M)” hazard level specified by GreenScreen®	3
	“High (H)” hazard level specified by GreenScreen®	4
Acute Human Toxicity (AT, ST-single, N-single, IrS, IrE)	“Low(L)” hazard level specified by GreenScreen®, strong evidence of negative studies, no hazard data, considered safe	1
	“Moderate (M)” hazard level specified by GreenScreen®	2
	“High (H)” hazard level specified by GreenScreen®	3
	“very High (vH)” hazard level specified by GreenScreen®	4
Chronic Human Toxicity (ST-repeated, N-repeated, SnS, SnR)	Strong evidence of negative studies, no hazard data, considered safe	1
	“Low (L)” hazard level specified by GreenScreen®	2
	“Moderate (M)” hazard level specified by GreenScreen®	3
	“High (H)” hazard level specified by GreenScreen®	4
Environmental Toxicity & Fate (AA, CA, P, B)	“very Low (vL) or Low(L)” hazard level specified by GreenScreen®, strong evidence of negative studies, no hazard data, considered safe	1
	“Moderate (M)” hazard level specified by GreenScreen®	2
	“High (H)” hazard level specified by GreenScreen®	3
	“very High (vH)” hazard level specified by GreenScreen®	4
Physical Hazards (Rx, F)	“Low(L)” hazard level specified by GreenScreen®, strong evidence of negative studies, no hazard data, considered safe	1
	“Moderate (M)” hazard level specified by GreenScreen®	2
	“High (H)” hazard level specified by GreenScreen®	3
	“very High (vH)” hazard level specified by GreenScreen®	4

2.2. Chemical Hazard Assessment

We used GreenScreen® for Safer Chemicals version 1.4 (GreenScreen®), developed by Clean Production Action, to assess chemicals based on their hazard endpoints, which allows for the conversion of the results into “benchmark scores” using transparent and systematic criteria (Clean Production Action, 2021a). GreenScreen® has become widely accepted and incorporated into various standards and programs that promote the minimization of toxic chemicals usage, such as Cradle-to-Cradle Certified®, EPEAT Registry, and TCO certified (Cradle to Cradle Products Innovation Institute, 2021; Global Electronics Council, 2021; Swedish Confederation of Professional Employees, 2021). GreenScreen® is also applied in many industrial sectors with its certification programs for the identification of products that may

Table 3
Scenarios used for data aggregation, uncertainty analysis and decision analysis.

Scenarios	Materials for evaluation						Description
	BM-1	BM-2	BM-3	BM-4	BM-U	BM-1, BM-2, and BM-3 converted to worst case	
Scenario 1	√	√	√	x	x	x	Baseline scenario – Considering primary materials with at least one hazard endpoint (BM-1, BM-2, and BM-3).
Scenario 2	√	√	√	√	x	x	Compensatory scenario – Considering primary materials with at least one hazard endpoint and those that are safer chemicals (BM-1, BM-2, BM-3, and BM-4).
Scenario 3	√	√	√	x	x	√	Worst-case scenario – Considering primary materials with at least one hazard endpoint (BM-1, BM-2, and BM-3), but for toxicity groups with data gaps, assume the worst possible score (s).
Scenario 4	√	√	√	x	√	x	Supplemental scenario – Considering primary materials with at least one hazard endpoint (BM-1, BM-2, and BM-3) and those with insufficient data (BM-U); the scores for BM-U primary materials are derived based on their associated reagents and intermediates.
Scenario 5	√	√	√	√	√	√	Integrated scenario – Considering all the primary materials used (BM-1, BM-2, BM-3, BM-4, and BM-U), and integrating the changes in Scenarios 2 – 4.

contain hazardous chemicals such as firefighting foam, textile chemicals, furniture & fabrics, and cleaners & degreasers in manufacturing (Clean Production Action, 2021b). GreenScreen® utilizes information on 20 hazard endpoints (Table 1), including endpoints related to human health, environmental toxicity and fate, and physical hazards. The selection and evaluation of these 20 hazard endpoints align with several national and international protocols such as the Globally Harmonized System of Classification and Labeling of Chemicals (GHS), the European Union's Registration, Evaluation, Authorization and Restriction of Chemicals (REACH), and the U.S. Environmental Protection Agency's Design for Environment Program (European Council, 1999; United Nations, 2019, United States Environmental Protection Agency, 2021b).

After completing the collection of toxicity information for the hazard endpoints, a decision logic is applied to generate one of five possible benchmark (BM) scores for each chemical: "Chemical of High Concern" (BM-1), "Use but Search for Safer Alternatives" (BM-2), "Use but Still Opportunity for Improvement" (BM-3), "Safer Chemical" (BM-4), and "Unspecified due to Insufficient Data" (BM-U). To conduct a GreenScreen® assessment, four steps specified in the guidance manual are followed, as described in detail in the SI Section S2.1. The GreenScreen® method requires a minimum amount of hazard endpoint data to assign a BM score, otherwise BM-U is assigned due to data gaps. In the current study, a BM-U score is assigned when no information can be found for the given substance. An official GreenScreen® assessment requires a third-party validation, which was not done here, therefore, these benchmark scores are currently designated as 'GreenScreen®-based'.

In this study, CHA is conducted and BM scores are assigned for each *primary material* in each battery using the GreenScreen® approach and the following toxicity data sources, which are further described in SI Section S2.1: GHS-Japan (National Institute of Technology and Evaluation, 2021), GESTIS (Institute for Occupational Safety and Health of the German Social Accident Insurance, 2021), European Chemicals Agency Registered Substances (ECHA CHEM) (European Chemicals Agency, 2021a), Hazardous Substances Data Bank (HSDB) (National Library of Medicine, 2021), EPI SUITE™ (United States Environmental Protection Agency, 2021c), EU SVHC List (European Chemicals Agency, 2021b) and The Endocrine Disruption Exchange (TEDX) List (The Endocrine Disruption Exchange, 2021). Several of the primary materials, however, are polymers and compounds for which no direct hazard and toxicity endpoint data exist. Thus, these primary materials are assigned benchmark scores of BM-U, which unfortunately isn't very useful when incorporating these into an overall assessment of the batteries themselves. A review of the literature indicates, however, that there is no established methodology for conducting CHA on such materials, so we have developed our own approach: to assess these polymers and compounds: we expanded the system boundary by evaluating the *reagents and intermediates* used to produce these BM-U materials. Both scenarios are considered later in this study. Based on the capability of Ecoinvent (Ecoinvent, 2021) for providing extensive data on up-stream chemicals, combined with a review of the pertinent literature and industry reports, these BM-U materials, as well as the '*active*' primary materials are thus expanded for further assessment. '*Active*' primary materials include *cathode* materials in LIBs and *electrolyte* materials in RFBs. It should be noted that the system boundary for reagents and intermediates is constrained to not include chemicals used for basic production activities such as mineral extraction and petroleum refining, as these would be common across the battery systems studied, providing little comparative insight among the battery types studied here. Details on the reagents and intermediates assessed in this study are provided in SI Section S3.

2.3. Hazard data quantification and aggregation - Scenario and decision analysis

The 'GreenScreen®-based' benchmark criteria allow us to assign a BM score for each primary material or reagent and intermediate assessed, which provides qualitative insight into the relative levels of hazard associated with the chemicals and batteries, as discussed in the Results section. However, the BM scores lack the resolution to be quantitatively aggregated for comparing these engineering products made from multiple components and materials. Moreover, the raw data upon which the BM scores are derived, which include 20 *hazard endpoints* combined into five *hazard groups* (see Table 1), are difficult to process mathematically. Therefore, we developed a modified assessment approach to quantitatively aggregate the hazard and toxicity data and to address uncertainties caused by the use of various data sources and the existence of data gaps. This approach applies the following steps, each of which is described below: (1) reclassifying *hazard endpoints* into *toxicity groups*, (2) harmonizing *hazard levels* and converting *qualitative hazard descriptions* into *quantitative hazard scores*, and (3) aggregating *hazard scores for materials*, both by hazard frequency and by material mass, into *single-value hazard scores for each battery product*. Ultimately, uncertainty is evaluated through the application of scenario analysis combined with decision analysis.

2.3.1. Refinement of hazard endpoints classification scheme

A framework for refining hazard endpoints into toxicity groups for quantification and high-resolution interpretation was explored by Faludi et al. (2016). We build on this approach, which evaluates the hazard endpoints directly (without implementing the benchmarking decision logic) and reclassifies them into six '*toxicity groups*'. In our scheme, the "Human Health Group I" category is divided into two toxicity groups as "Carcinogenic and Mutagenic Toxicity" (C/M), and "Reproductive, Developmental, and Endocrine Toxicity" (R/D/E). The previous "Human Health Group II" category is now modified as the "Acute Human Toxicity" group, while the "Human Health Group II* " category is now changed to the "Chronic Human Toxicity" group in

Table 4
The mass percentages (Majeau-Bettez et al., 2011; Notter et al., 2010) and GreenScreen®-based benchmark scores for the major components and primary materials used in the three lithium-ion batteries: lithium iron phosphate (LFP), lithium nickel cobalt manganese hydroxide (NCM), and lithium manganese oxide (LMO).

Detail specification	**LFP**		**NCM**		**LMO**	
Mass per kWh capacity	7.86 kg/kWh		6.15 kg/kWh		6.80 kg/kWh	
			Cell stack component			
Cathode	⬜ Lithium iron phosphate	31.2%	🟥 Lithium nickel cobalt manganese hydroxide	29.3%	⬜ Lithium manganese oxide	20.3%
	🟥 Carbon black	1.8%	🟥 Carbon black	1.7%	🟥 Carbon black	0.9%
	🟨 Polytetrafluoroethylene	2.9%	🟨 Polytetrafluoroethylene	2.7%	🟧 Latex	0.3%
	🟥 N-methyl-2-pyrrolidone	10.0%	🟥 N-methyl-2-pyrrolidone	9.4%		
Cathode substrate	🟧 Aluminum	5.2%	🟧 Aluminum	5.2%	🟧 Aluminum	12.8%
Anode	🟩 Graphite	11.0%	🟩 Graphite	13.0%	🟩 Graphite	19.7%
	🟨 Polytetrafluoroethylene	0.6%	🟨 Polytetrafluoroethylene	0.7%	🟥 Carbon black	0.6%
	🟥 N-methyl-2-pyrrolidone	3.2%	🟥 N-methyl-2-pyrrolidone	3.8%	🟧 Latex	0.8%
Anode substrate	🟧 Copper	12.0%	🟧 Copper	12.0%	🟧 Copper	20.9%
Electrolytes	🟧 Lithium Hexafluorophosphate	2.1%	🟧 Lithium Hexafluorophosphate	2.1%	🟧 Lithium Hexafluorophosphate	2.1%
	🟧 Ethylene carbonate	15.3%	🟧 Ethylene carbonate	15.3%	🟧 Ethylene carbonate	15.7%
Separator	🟧 Polyethylene	2.4%	🟧 Polyethylene	2.4%	🟥 Silica sand	1.2%
					🟥 Phthalic anhydride	1.5%
					🟥 Acetone	0.1%
	🟧 Polypropylene	2.4%	🟧 Polypropylene	2.4%	🟧 Polyethylene	1.9%
					🟨 Hexafluoroethane	0.1%
					⬜ Polyvinyl fluoride	1.0%

BM-1 🟧 **BM-2** 🟨 **BM-3** 🟩 **BM-4** ⬜ **BM-U**

Color scheme: red = BM-1, orange = BM-2, yellow = BM-3, green = BM-4, and gray = BM-U.

order to emphasize the single or repeated exposure, respectively. The "Environmental Toxicity & Fate" and "Physical Hazards" groups remain consistent with the classifications used in the GreenScreen® framework. While similar to the method used by Faludi et al. (2016), the above scheme used in the current study combines the environmental toxicity and fate, and the physical hazards are still under consideration in our case.

2.3.2. *Harmonization of hazard levels and conversion to quantitative hazard scores*

Another necessary adjustment is to harmonize the refined hazard categories and convert them to numerical scores. The GreenScreen® framework uses hazard levels for each hazard endpoint categorized as "very Low" (vL), "Low" (L), "Moderate" (M), "High" (H), and "very High" (vH), in accordance with the GHS labeling system (United Nations, 2019). In the current study, we again build on the approach developed by Faludi et al. (2016) to convert the hazard levels of very low (vL) to very high (vH) into hazard score values from 1-to-4, with the 1 denoting no hazard found in current toxicity data sources and 4 representing the highest hazard level, *thus low hazard score values are preferred*. We also harmonize the different ranges for hazard levels of different hazard endpoints (e.g., carcinogenicity (C) only has hazard levels ranging from low (L) to high (H), while persistence (P) includes hazard levels from very low (vL) to very high (vH)), as detailed in Table 2. The *hazard scores for each material* are calculated as the average of the hazard scores for the six toxicity groups for that substance. Lastly, if the hazard scores for the individual hazard endpoints within a given toxicity group do not agree, we use the highest hazard score within that toxicity group. For instance, if a chemical is identified to receive a hazard score of 3 for carcinogenicity (C) and a hazard score of 4 for

Table 5
The mass percentages (He et al., 2020) and GreenScreen®-based benchmark scores for the major components and primary materials used in the three flow batteries – vanadium redox flow battery (VRFB), zinc bromine flow battery (ZBFB), and all iron flow battery (IFB).

Detail specifications	**VRFB**		**ZBFB**		**IFB**	
Mass per kWh capacity	18.68 kg/kWh		14.40 kg/kWh		32.52 kg/kWh	
Cell stack component						
Bipolar plate	Graphite	1.6%	Titanium	7.8%	Graphite	6.0%
	Polyethylene	0.1%	Polyethylene	8.3%	Bisphenol-A epoxy-based vinyl ester resin	1.1%
Electrode	Carbon fiber felt	0.2%	/[a]		Carbon fiber felt	0.1%
Membrane	Nafion®	0.1%	/[a]		Polyethylene	0.1%
Cell frame	E-glass fiber	0.4%	Polyethylene	6.1%	E-glass fiber	5.3%
	Polypropylene	1.8%			Polyester resin	3.6%
Current collector	Copper	1.6%	Titanium	1.7%	Aluminum	1.0%
Electrolyte storage component						
Electrolytes	Vanadium pentoxide	58.9%	Zinc bromide	36.7%	Iron (II) chloride	32.3%
	Hydrochloric acid	16.1%			Potassium chloride	35.5%
	Sulfuric acid	5.4%	Bromine	31.1%	Manganese dioxide	2.2%
					Hydrochloric acid	1.9%
Tank	Polyethylene	13.8%	Polyethylene	5.6%	Isophthalic polyester	9.2%
Pipe	Polyethylene	0.2%	Polyethylene	1.7%	Polyvinylchloride	1.8%

[a] No extra electrode and membrane materials required for ZBFB

BM-1 🟧 **BM-2** 🟨 **BM-3** 🟩 **BM-4** ▨ **BM-U**

Color scheme: red = BM-1, orange = BM-2, yellow = BM-3, green = BM-4, and gray = BM-U.

mutagenicity (M), then the hazard score for the C/M toxicity group would be 4.

2.3.3. *Aggregation of hazard scores*

Another innovation was necessary to address the challenge of determining how to aggregate hazard scores of individual materials to generate a *single-value hazard score* for the *battery*. Little guidance exists in the literature, except a previous study on thin film solar panels in which the frequency of toxicity scores was tracked (Eisenberg et al., 2013) and another study on utility meters in which results were compared based on both frequency and mass fraction (Lam et al., 2013). The dilemma here is whether a toxic substance is of equal concern simply by being present, or can its impact be diluted if offset by less toxic substances present in higher concentrations. We model our approach off the Lam et al. (2013) study, providing results for both *'number-based' and 'mass-based' aggregation*. 'Number-based' aggregation assumes each material is of equal importance and mathematically averages the hazard scores of all materials to determine the single-value hazard score for the battery. 'Mass-based' aggregation assumes the relative importance of a given material is determined by its mass fraction, and the single-value hazard score for the battery is calculated accordingly.

2.3.4. *Scenario analysis*

To account for data gaps, data uncertainties and assumptions used for the quantification and aggregation, it is important to conduct scenario analysis before drawing robust conclusions regarding the relative single-value hazard scores for the six batteries of interest. Five scenarios are included in this analysis, as summarized in Table 3. Scenario 1 is the baseline scenario and includes in the aggregated single-value hazard scores only materials that were assigned benchmark scores of BM-1, BM-2 or BM-3. Scenario 2 is considered the 'compensatory scenario' because it includes materials assigned benchmark scores of BM-4 ('Safer Chemical'), which consequently have hazard scores of 1.0 for all six toxicity groups. These low hazard scores, which are based on fairly limited data, dilute the higher scores of the BM-1 to BM-3 materials, and therefore compensate for the other more hazardous materials in the battery. For Scenario 3, data gaps within a given toxicity group are addressed; specifically, missing data, which may be determined at a future date through further testing and modeling, are assumed to have the highest (worst) hazard score of 4.0. For Scenario 4, the materials with benchmark scores of BM-U are included; hazard scores for these materials are determined using the expanded system boundary approach assessing reagents and intermediates, as described above. Finally, Scenario 5

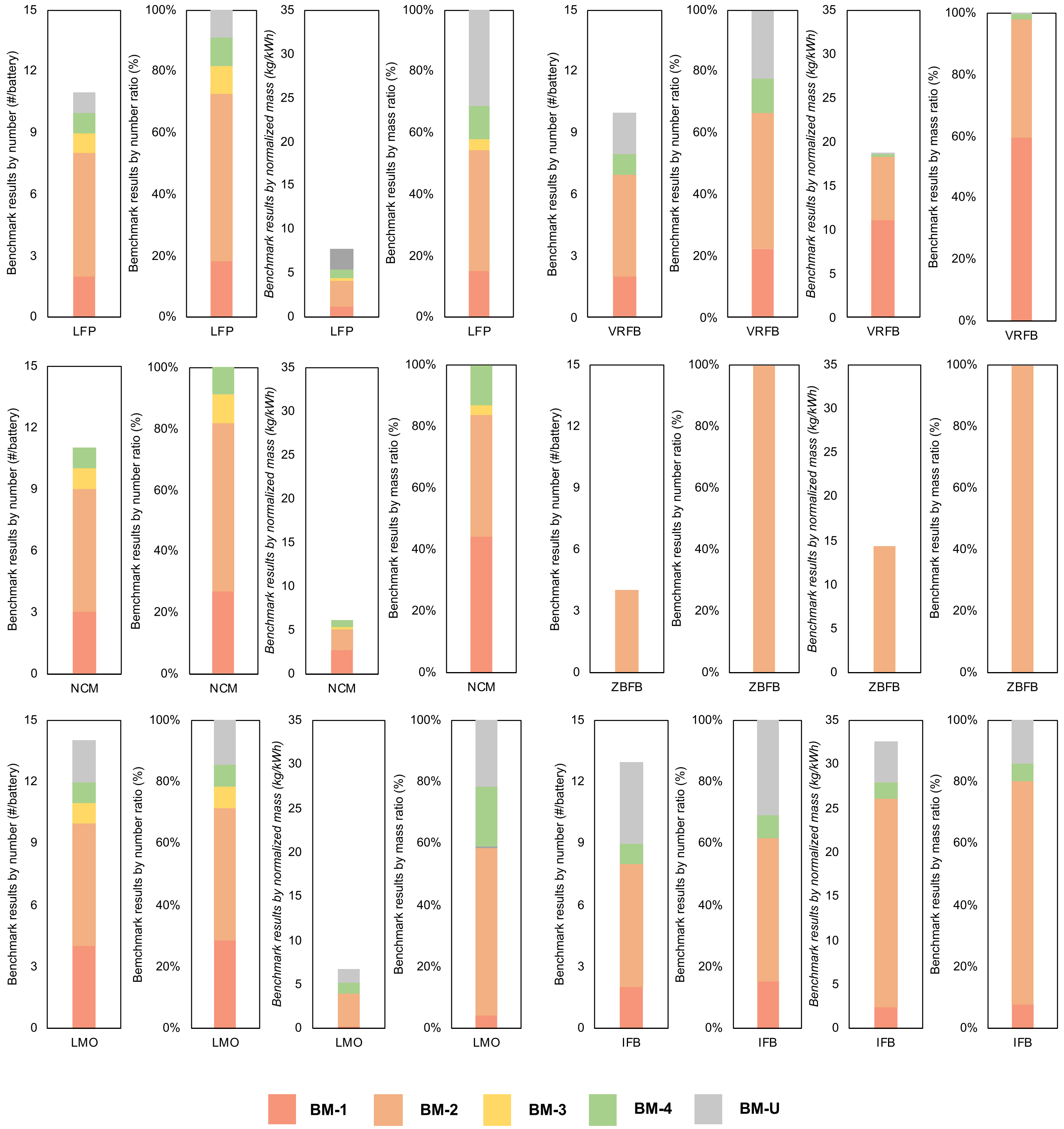


Fig. 2. The GreenScreen®-based benchmark score results according to number of materials (#/battery), number ratio (%), normalized mass (kg/kWh energy capacity) and mass ratio (%), for the primary materials in the three lithium-ion batteries (LFP = lithium iron phosphate, NCM = lithium nickel cobalt manganese hydroxide, and LMO = lithium manganese oxide) and the three redox flow batteries (VRFB = vanadium redox flow battery, ZBFB = zinc-bromine flow battery, IFB = all-iron flow battery). BM-1: chemical of high concern; BM-2: use but search for safer alternatives; BM-3: use but still opportunity for improvement; BM-4: safer chemical; and BM-U: unspecified due to insufficient data.

simultaneously integrates all of the modifications incorporated into Scenarios 2 through 4.

2.3.5. Multi-criteria decision analysis

The data aggregation approach described above successfully converts an extensive data set on 20 hazard endpoints for numerous materials into a single-value hazard score for each battery product according to 6 toxicity groups. However, it is important to ascertain whether each toxicity group should be considered of equal importance, or weighted differently. There are few studies in the literature that explore the weighting scheme for distinct hazard endpoints (Hu et al., 2019; Malloy et al., 2017), yet there is no consensus, and strategies tend to depend on various stakeholder preferences. A recent OECD report provided an aggregation scheme that used the sum of all the hazard

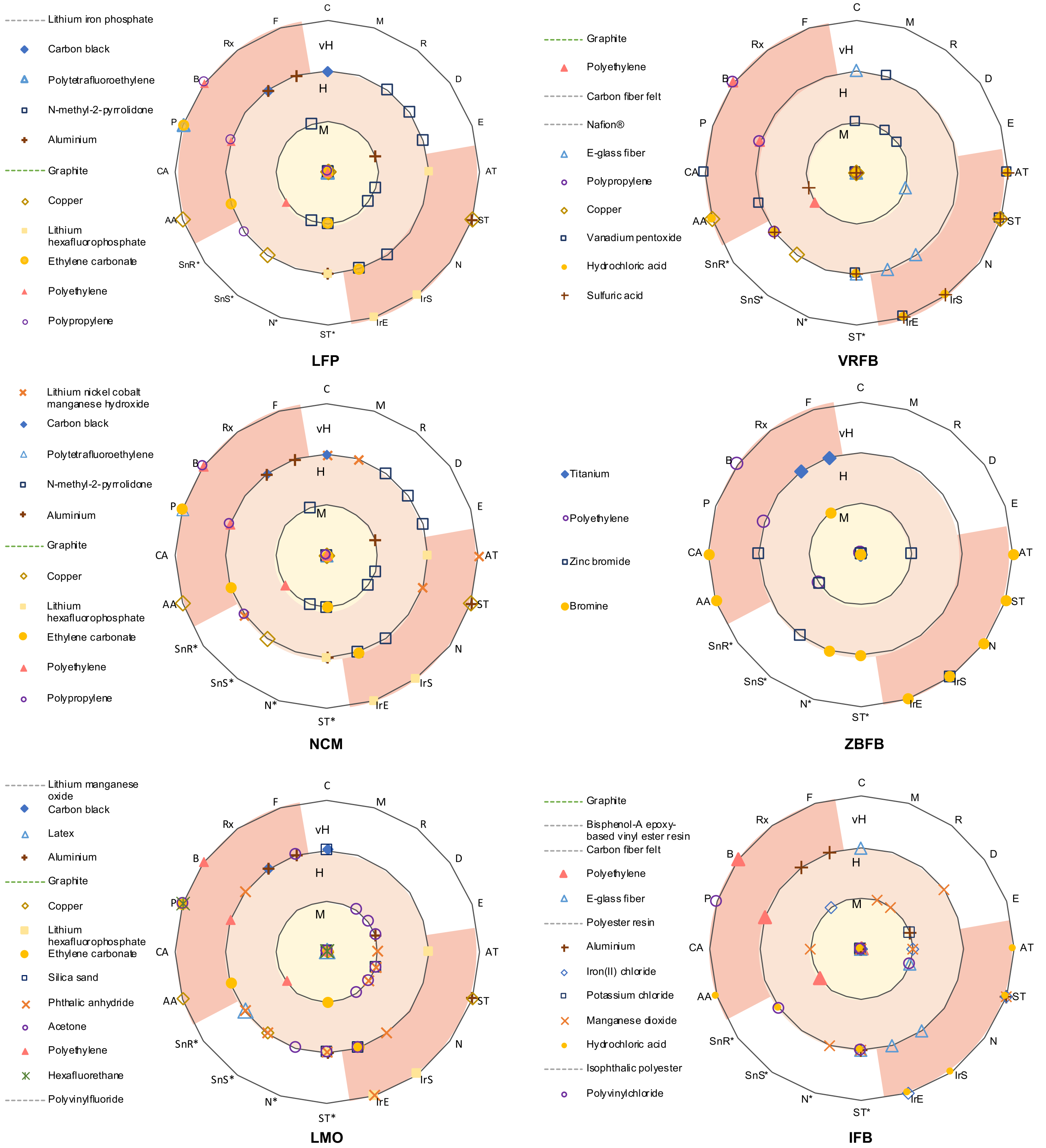


Fig. 3. The distribution of the 20 hazard endpoints (see Table 1) triggered by the primary materials used in each of the six batteries. The primary materials identified as BM-4 and BM-U materials are not included in the plots, but they are listed in the legend with green and gray dashed lines, respectively. *Values closer to the center of the plots are preferred.*

triggers, which implied equal weighting for each hazard endpoint (OECD, 2021). Faludi et al. proposed a different set of weighting factors for scenario analysis, but ultimately utilized expert judgement (Faludi et al., 2016). We have chosen to apply MCDA, because it is a quantitative and systematic approach that assesses multiple criteria problems and can be used to explore strategies for dealing with weighting factors. In our case, the six battery products are identified as the '*alternatives*' selected for comparison, the '*criteria*' for evaluation are the 6 toxicity groups, and stochastic multicriteria acceptability analysis (SMAA) is the MCDA method chosen for analysis (Lahdelma et al., 1998; Lahdelma and Salminen, 2001).

SMAA is chosen for this analysis because it deals with multi-objective problems by exploring the weight space to characterize the '*performance*' of selected alternatives (i.e., referred to hereafter as the '*hazard ranking*'

Table 6
The GreenScreen®-based benchmark scores for the reagents and intermediates used to produce active primary materials in the six batteries.

Primary materials	BM-1	BM-2	BM-3	BM-4	BM-U	Total	Materials Usage
Lithium iron phosphate BM-U	2	7	0	0	0	9	LFP cathode active material
Lithium nickel cobalt manganese hydroxide BM-1	5	6	0	0	1	12	NCM cathode active material
Lithium manganese oxide BM-U	2	4	0	2	0	8	LMO cathode active material
Vanadium pentoxide BM-1	0	3	0	0	0	3	VRFB electrolyte active material
Zinc Bromide BM-2	0	4	0	0	0	4	ZBFB electrolyte active material

Table 7
The GreenScreen®-based benchmark scores for the reagents and intermediates used to produce select primary materials receiving BM-U scores.

Primary materials	BM-1	BM-2	BM-3	BM-4	BM-U	Total	Materials Usage
Polyvinyl fluoride BM-U	1	5	0	0	0	6	LMO separator
Carbon fiber felt BM-U	5	5	0	0	1	11	VRFB and IFB electrode
Nafion® BM-U	3	8	0	1	2	14	VRFB membrane
Bisphenol-A epoxy-based vinyl ester resin BM-U	7	10	1	1	0	19	IFB bipolar plate filler
Polyester resin BM-U	8	10	0	1	0	19	IFB cell frame filler
Isophthalic polyester BM-U	10	6	0	1	0	17	IFB tank

of each battery product relative to the others) (Lahdelma et al., 1998; Lahdelma and Salminen, 2001). Hence, the SMAA differs from other MCDA methodologies, as the '*alternative's performance*' is not a specified utility score on each criterion but a stochastic variable, which is deemed to be an inverse approach that allows us to not provide predetermined weighting factors. In SMAA, the collection of all the possible weighting factors, which is also called the '*acceptability index*', is distributed to each alternative like a probabilistic distribution. For a simple illustration, the deterministic case is explained through Eqs. (1) – (6). As seen in Eq. (1), the basic idea of MCDA is to determine the total utility U($\boldsymbol{a}$) of alternative $\boldsymbol{a}$ in terms of its '*performance*' on criterion i as u($\mathbf{a}_i$), and $\boldsymbol{w}_i$ is the weighting factor for criterion i while $\boldsymbol{w}_i > 0$ and $\Sigma \boldsymbol{w}_i = 1$.

$$\mathrm{U}(\mathbf{a}) = \Sigma\, \mathbf{w}_i \times \mathrm{u}(\mathbf{a}_i) \tag{1}$$

For SMAA, the weighting factor combinations $\boldsymbol{w}$ that yield an overall utility for alternative $\boldsymbol{a}$ greater than or equal to that for any other alternative is introduced as:

$$U(\boldsymbol{a}) \geq U(\boldsymbol{b}),\ \forall \boldsymbol{b} \tag{2}$$

$$\Sigma\, \boldsymbol{w}_i \times u(\boldsymbol{a}_i) \geq \Sigma\, \boldsymbol{w}_i \times u(\boldsymbol{b}_i) \tag{3}$$

In the next step, a set of weighting factor combinations $\boldsymbol{w}$ that will satisfy Eq. (2) are referred to as favorable weighting vectors W_a. As mentioned, an '*acceptability index*' A_i can now be determined as the volume of the favorable weighting vectors divided by the volume of total feasible weighting vectors $\boldsymbol{W}$ where:

$$W = \{\, \boldsymbol{w} \in R^n\!: \boldsymbol{w}_i > 0 \text{ and } \Sigma\, \boldsymbol{w}_i = 1 \,\} \tag{4}$$

$$W_a = \{\ \boldsymbol{w} \in W\!: U(\boldsymbol{a}) \geq U(\boldsymbol{b}),\ \forall \boldsymbol{b}\ \} \tag{5}$$

$$A_i = vol(W_a)/vol(W) \tag{6}$$

In our case, the specific method applied is SMAA-2 (Lahdelma and Salminen, 2001), which is an extension to SMAA, as it calculates the *acceptability index for each alternative battery product*, to be ranked not only first, but from best to worst, and the JSMAA software is used to construct the SMAA-2 model (Tervonen, 2014). The evaluation of both 'number-based' and 'mass-based' aggregation, and the different scenarios, are also applied in the MCDA.

3. Results and discussion

3.1. Chemical hazard assessment results

GreenScreen®-based benchmark scores were determined for each

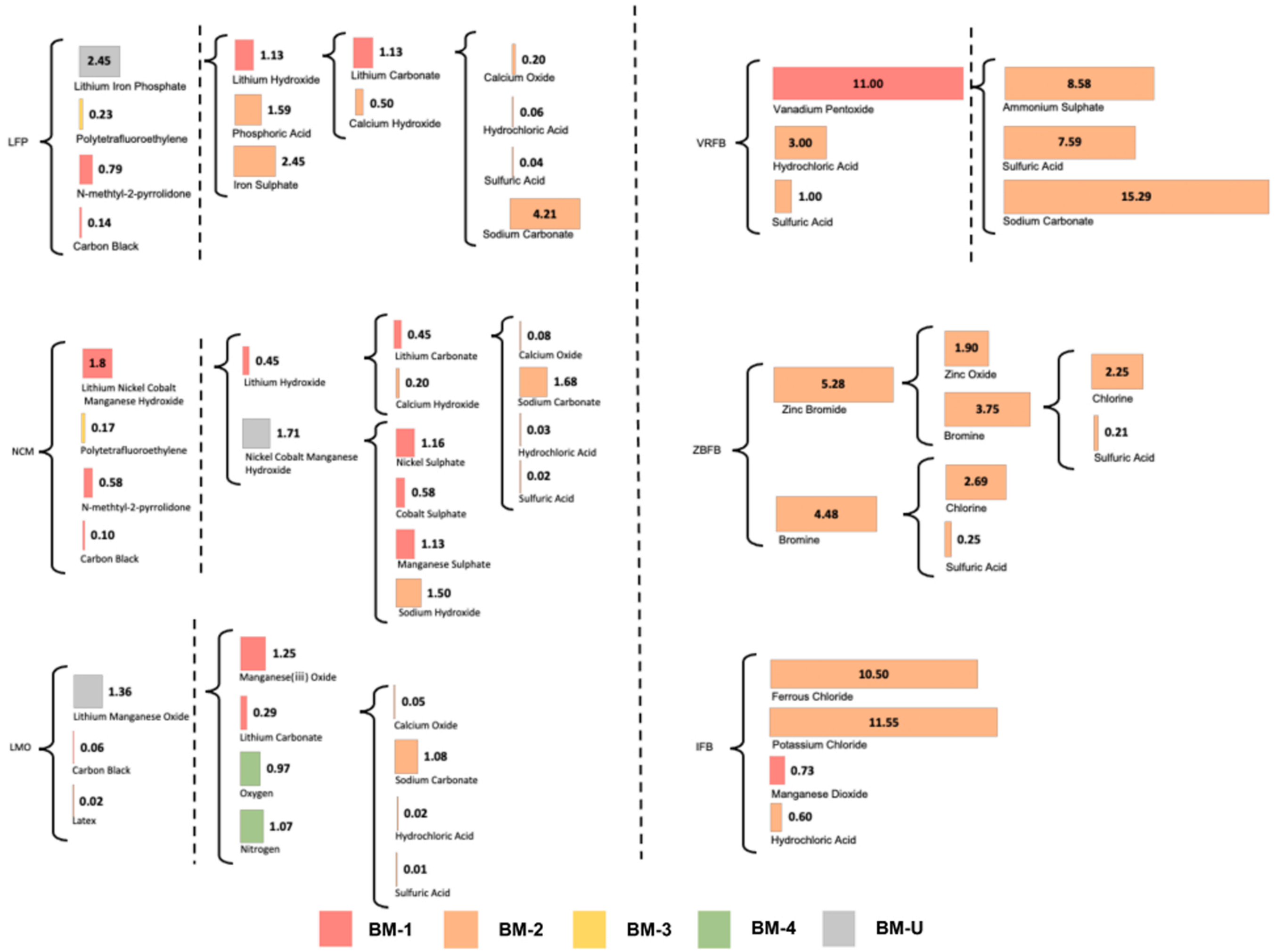


Fig. 4. The GreenScreen®-based benchmark score results for the reagents and intermediates used to produce active primary materials in each battery product. BM-1: chemical of high concern; BM-2: use but search for safer alternatives; BM-3: use but still opportunity for improvement; BM- 4: safer chemical; and BM-U: unspecified due to insufficient data.

primary material used in each of the six battery products, as shown in Tables 4 and 5, for LIBs and RFBs, respectively. The color coding in these tables helps to visualize the distribution of materials across the possible benchmark scores, noting the distribution of number of materials in each BM category as BM-2 > BM1 > BM-3 > BM-4, plus quite a few BM-U materials (BM-4 materials are preferred to BM-3, etc., and BM-U is 'unspecified' due to insufficient data). To further visualize the distribution in benchmark scores for each battery product, the data presented in Fig. 2 show the following metrics: benchmark results by absolute number, benchmark results by number ratio, benchmark results by normalized mass and benchmark results by mass ratio, with the latter two normalized per kWh capacity. Notably, BM-2 materials dominate regardless of the metric, BM-U materials are used in significant numbers and mass fractions, and the use of BM-1 materials is not insignificant. It is also noted that vanadium pentoxide, a BM-1 material used as the electrolyte in VRFBs, corresponds to a high mass faction (~60 %) in the VRFB system. Among the battery types, ZBFB is the only one that does not consist of any BM-1 or BM-U materials; and overall, the mass per kWh for LIBs is notably less than that for RFBs.

The hazard endpoints that lead to the benchmark scores can also be visually compared, as shown in Fig. 3, *noting that data points closer to the center of the spider plots are preferred*. These plots highlight the variable scales used for different endpoints, as described in Section 2.3.2, making direct comparison, even visually, a bit difficult. Yet a few key findings can be derived from the plotted data. Almost all of the 20 hazard endpoints are triggered, once or several times, by the primary materials used in each battery, yet among the six batteries, ZBFB is the only battery that shows no materials triggering any Human Health Group I hazard endpoints, and VRFB is the only one with none of its materials triggering Physical Hazards. Looking closer at these materials in Fig. 3 along with their GreenScreen®-based benchmark scores in Tables 4 and 5, we observe that several materials, including carbon black, N-methy-2-pyrroidone, lithium nickel cobalt manganese hydroxide, silica sand, phthalic anhydride, acetone, e-glass fiber, vanadium pentoxide, and manganese dioxide, are chemicals of high concern (BM-1), and their use should be replaced by safer alternatives. For example, N-methyl-2-pyrroidone, used in the LFP and NCM, should be avoided as it triggers high hazard on three of the five hazard endpoints in Human Health Group I: reproductive toxicity (R), developmental toxicity (D), and endocrine activity (E).

We can also expand the system boundary for the chemical hazard assessment to include reagents and intermediates, especially for the BM-U materials, as described above (in Section 2.2), and for the active primary materials, which are critical to the functionality of the batteries. These results are presented visually in Tables 6 and 7, and Figs. 4 and 5. Clearly, there is a large number of reagents and intermediates that are used to produce the primary materials, regardless of battery product, many of which exhibit high toxicity (BM-1 and BM-2 scores).

The data presented in Tables 4–6 and Fig. 4 show that two of the three LIB active primary materials, i.e., for the cathode, lithium iron

Fig. 5. The GreenScreen®-based benchmark score results for the reagents and intermediates used to produce select primary materials that identified as BM-U materials. BM-1: chemical of high concern; BM-2: use but search for safer alternatives; BM-3: use but still opportunity for improvement; BM-4: safer chemical; and BM-U: unspecified due to insufficient data.

phosphate and lithium manganese oxide, are identified to be BM-U materials, and the cathode material for the third LIB battery type, lithium nickel cobalt manganese, is identified to be BM-1. Note that the two BM-U materials represent approximately 30 wt % and 20 wt % of their respective batteries. For the active primary materials, i.e., the electrolytes, in RFBs, vanadium pentoxide is classified as BM-1, whereas zinc bromide, bromine, ferrous chloride, potassium chloride are all BM-2; the only other BM-1 material is manganese dioxide in IFB's (see Table 5 and Fig. 4). Note that vanadium pentoxide represents approximately 60 wt % of the VRFB. Looking next at the reagents and intermediates used to produce these active primary materials (see Table 6 and Fig. 4), LIBs required the use of more reagents and intermediates compared to RFBs, and many of them are identified to be BM-1 materials, such as lithium, nickel and manganese compounds. The manufacturing chains for these materials in LIBs are longer than for RFBs because the production of these active cathode materials require more complex synthesis procedures. All of the reagents and intermediates evaluated for RFBs are identified to be BM-2 materials. Lastly, the data in Table 7 and Fig. 5 summarize the benchmark scores for the reagent and intermediates needed to make the remaining BM-U primary materials. A total of 81 chemicals are evaluated, 34 of which (42 %) are BM-1 materials and 39 of which (48 %) are BM-2 materials, indicating that more than 90 % of these substances (73 out of 81) are chemicals of concern. Clearly there is a need to find safer alternative substances for producing the primary materials used in all six battery products.

Overall, the chemical hazard assessment and benchmark scores show that a wide range of toxic and hazardous materials are used to produce the six battery types. Many of these are BM-1 materials due to their potential carcinogenicity, mutagenicity and genotoxicity, reproductive and developmental toxicity, and endocrine disruptive activity. Efforts should be made to reduce the use of these materials, especially if they are not key to the active functionality of the battery. Beyond these insights, however, direct quantitative comparisons among the battery products are difficult when using GreenScreen®-based benchmark scores alone, as described previously. Therefore, we continue by presenting the quantitative data aggregation and scenario analysis, below.

3.2. Data aggregation and scenario analysis results

The single-value hazard scores for each battery, as derived through data aggregation, are presented in Figs. 6 and 7. Results for each of the

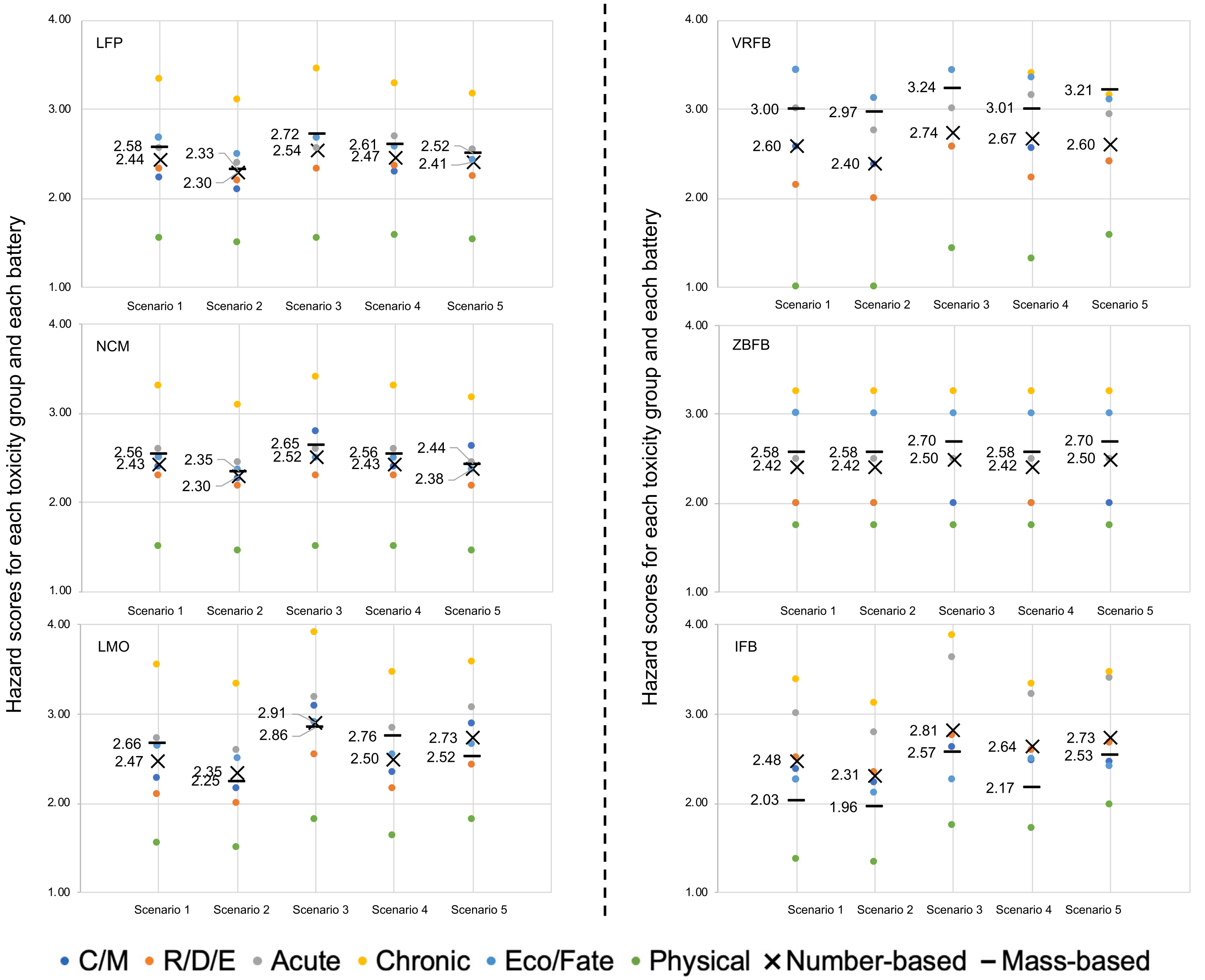


Fig. 6. The data aggregation results represented by single-value hazard scores for each toxicity group (shown as colored dots) and each battery product. The five scenarios are defined in Table 3. Both 'number-based' and 'mass-based' aggregation results (shown as black and gray bars, respectively) for each battery product are also provided. *Lower values are preferred.*

five scenarios are presented, for both 'number-based' and 'mass-based' scenarios. The hazard scores for each toxicity group are also presented to inform interpretation of the results. *Recall that the higher the score, the higher the toxicity or hazard, thus low values are preferred.*

When comparing the various scenarios, regardless of battery type, Scenario 2 lowers the single-value hazard score for each battery compared to Scenario 1 due to the compensatory effect of including the safer chemicals (i.e., BM-4 materials) used in each battery product. In contrast, Scenario 3 and Scenario 4 increase the single-value scores due to the inclusion of more hazard triggers or toxic reagents and intermediates into the analysis, especially for Scenario 3 which assumes the worst-case for missing hazard endpoint data. The results for Scenario 5 are more complex because of the trade-offs among Scenarios 2, 3 and 4.

From these quantitative representations, a few findings are highlighted. Among the different battery products, the VRFB tends to receive an overall higher single-value hazard score relative to the other five batteries especially for the 'mass-based' scenarios, and there are noticeable variations between the 'mass-based' scores and the 'number-based' scores in the five scenarios. The reason for this is the VRFB electrolyte materials, specifically vanadium pentoxide, which represents a high mass ratio and exhibits higher toxicity than other materials used in the battery products. For ZBFB, the single-value hazard score is insensitive to Scenarios 2 and 4 as no BM-4 and BM-U materials are used in this battery product. Also, the increase on Scenario 3 is slight, which indicates less uncertainty due to the hazard endpoints for the currently utilized materials in this battery product. Thus, the results for ZBFB are the most consistent among the batteries investigated. The single-value hazard scores for the LFP and NCM batteries also do not vary much across the different scenarios, and their scores are consistently close in value since they share many of the same materials. The single-value hazard score for IFB is more variable across the different scenarios. Also, the 'mass-based" score for IFB is lower than the 'number-based' score, which makes it different from the other batteries. This occurs because the IFB electrolyte material corresponds to a high mass ratio, similar to VRFB and ZBFB, but corresponds to lower toxicity. For LMO batteries, the single-value hazard score is also variable across the different scenarios, and the relative values for the 'number-based' and the 'mass-based' results also depend on the scenario. This occurs because of the inclusion in Scenario 2 of graphite, which is a BM-4 material that corresponds to 20 % of the total mass, and the use of worst-case hazard scores to fill data gaps in Scenario 3.

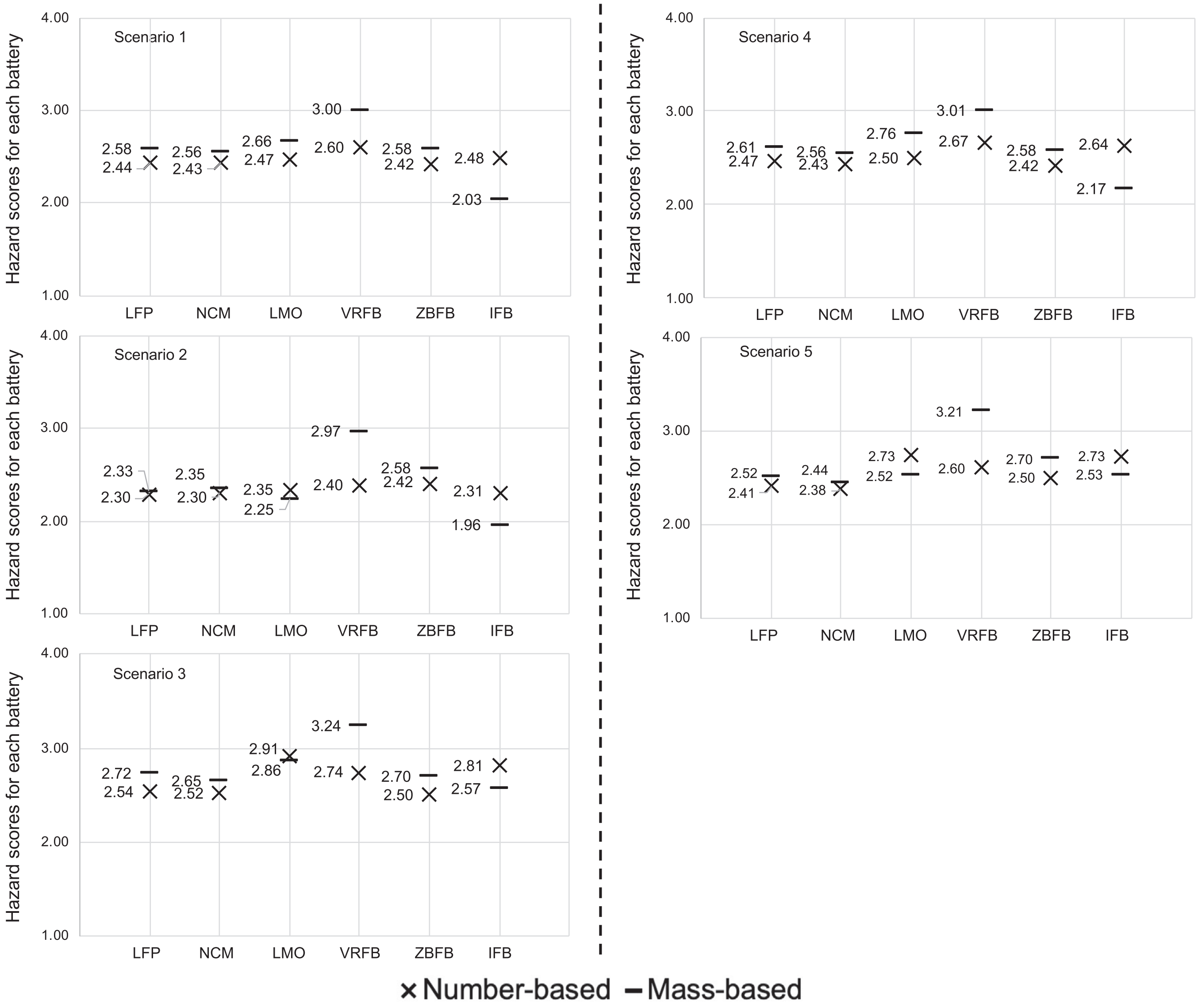


Fig. 7. The data aggregation results represented by single-value hazard scores for each battery product and each scenario (as defined in Table 3). Both ‘number-based’ and ‘mass-based’aggregation results (shown as black and gray bars, respectively) for each battery product are also provided. *Lower values are preferred.*

Looking closer at the toxicity group hazard scores, the “Chronic Human Toxicity” group usually shows the highest score, while the “Physical Hazards” group shows the lowest score, no matter what type of battery is chosen. VRFB and ZBFB exhibit higher toxicity in the “Environmental Toxicity & Fate” group compared to the other batteries, and IFB exhibits the highest toxicity in the “Acute Human Toxicity” group, followed by VRFB and LMO. The ZBFB has the lowest score in both the “Carcinogenic and Mutagenic Toxicity” group and “Reproductive, Developmental, and Endocrine Toxicity” group, as no substances used in ZBFB exhibit these hazard endpoints.

The aggregation into a single-value hazard score makes the comparison of different batteries more intuitive then with the chemical hazard assessment, as there is no longer a need to consider various battery materials and multiple hazard endpoints. However, the high-level aggregation leads to fairly similar results for each battery, which neutralizes efforts to make clear and concise decisions regarding which one is less toxic relative to the others. The trade-offs exhibited by the single-value hazard score indicate that the high-level aggregation is not sufficient for disclosing the complete hazard profile of materials or products for evaluation. Instead, it is an important intermediate step that allows us to apply decision analysis, which is more mathematically and logically rigorous. Thus, we continue with the decision analysis approach by introducing the multi-criteria decision analysis method, SMAA.

3.3. *Multi-criteria decision analysis results*

In Fig. 8, the SMAA results for the six batteries and the five scenarios are presented, with the ‘number-based’ and ‘mass-based’ results provided separately. In an effort to address the weighting factors for each toxicity group and the similar values for the single-value hazard scores, SMAA is applied to derive the acceptability index (0–100 %, vertical axis in Fig. 8) for the six batteries, which is equivalent to the probability of each battery being selected for a given hazard ranking from first to last (1st to 6th, horizontal axis in Fig. 8) relative to the other batteries. The calculated acceptability index values for each battery are provided in Table S7 in the SI. *Note that in the SMAA modeling, the inverse normalization is used, thus, the higher the acceptability index, the lower the average toxicity, which is preferred.* For example, in the baseline scenario (Scenario 1), the hazard ranking of ZBFB is among the highest in the ‘number-based’ scenario, as the acceptability index for ZBFB being ranked 1st is 53 %, and no other batteries have an acceptability index

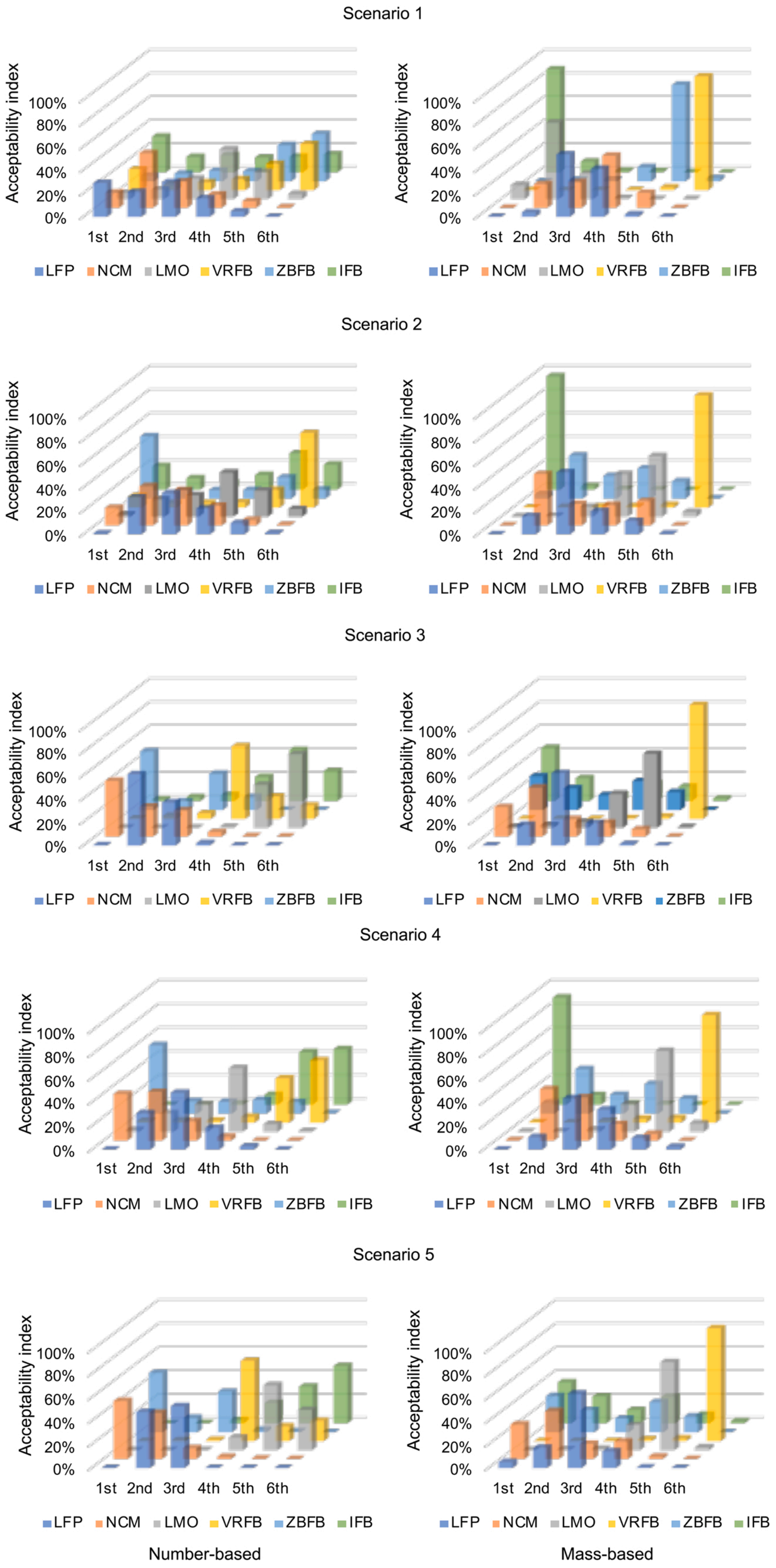


Fig. 8. The stochastic multicriteria acceptability analysis (SMAA) results based on the hazard score for each toxicity group for the six batteries in terms of 'number-based' (left panel) and 'mass-based'(right panel) aggregations, shown for each of the five scenarios (as defined in Table 3). To assist with interpreting the results, the numerical values for the acceptability index are provided in Table S7 in the SI. *The inverse normalization is taken, therefore, higher scores are preferred.*

over 20 % to be ranked 1st. IFB ranks the highest in the 'mass-based' scenario, as the acceptability index is as high as 96 % to be ranked 1st. In contrast, the hazard ranking of IFB in the 'number-based' scenario is the second lowest, given its acceptability index is more than 50 % for the 5th and the 6th rankings. VRFB ranks the lowest in both the 'number-based' and 'mass-based' scenarios, as noted by its highest acceptability index to be ranked 6th (63 % and 95 %, respectively). The three LIBs consistently rank in the middle. LMO is among the lowest with its peak acceptability index values of 37 % to be ranked 4th and 22 % to be ranked 5th in the 'number-based' scenario, and 36 % to be ranked 4th and 51 % to be ranked 5th in the 'mass-based' scenario. Interestingly, given the many similar materials used in the LFP and NCM batteries, it is not surprising that the acceptability index values for these two batteries are close, however, the hazard ranking of LFP in the 'mass-based' scenario (15 % for 2nd and 53 % for 3rd) is clearly one position lower than that of NCM (44 % for 2nd and 18 % for 3rd), which is attributed to the different mass ratio of toxic materials used (see Table 4) due to the different configurations of these battery products.

Although the SMAA results still exhibit some variations among the results when comparing across different scenarios, there are some clear distinctions among the hazard rankings associated with certain battery products regardless of scenario. For example, VRFB and LMO are among the lowest ranked in almost all the scenarios. IFB ranks highest in most of the 'mass-based' scenarios, but not for the 'number-based' scenarios. In contrast, ZBFB is ranked high in most of the 'number-based' scenarios, but not the 'mass-based' scenarios. Interestingly, the hazard rankings for NCM and LFP are consistently among the top 4 in all of the scenarios, especially in Scenario 5, which integrates all of the changes incorporated in Scenarios 2–4. This is primarily because NCM and LFP exhibit fewer uncertainties, as incorporated in Scenario 3, and the variations brought by Scenario 3 are the most prominent when considering the trade-offs of those different scenarios. Of course, it is important to note here that the relative rankings of these batteries are dependent on the materials evaluated in this assessment case study. The materials used in the batteries, as well as the intermediates and reagents, can change over time, affecting not only battery functionality and performance, but also the CHA results and consequential ranking relative to alternatives. The framework developed here can be easily updated to evaluate such changes, and especially to determine if the changes lead to relative improvements (i.e., reduced use of high-hazard materials), which is a primary motivation for this work.

4. Conclusions

The assessment on toxicity hazards for six batteries provides important guidance on promoting safer alternative materials for batteries production with the materials of high concern and specific hazard endpoints highlighted. From the CHA results, the hazard ranking of LIBs and RFBs vary considerably due to the different choices of materials and design of the battery structure and components, as well as the functional parameters such as energy density. In general, the LIBs use a greater number of highly hazardous materials, whereas the RFBs use a higher mass of highly hazardous materials. Regarding specific materials used in these batteries, of highest concern are the materials that are key to functionality, yet are assessed as either BM-1 or BM-U. For example, vanadium pentoxide and lithium nickel cobalt manganese hydroxide used in VRFB and NCM, respectively, are material of high concern triggering high hazard in Human Health Group I hazard endpoints. Furthermore, lithium iron phosphate and lithium manganese oxide used in LFP and LMO, respectively, have no hazard endpoints identified due to lack of toxicity data. The expansion of the system boundary to include reagents and intermediates highlights additional materials of concern. Consequently, as battery technology is developing rapidly, efforts should be made to evaluate hazard early in the design phase to avoid the need for costly substitutions after products and manufacturing details are fully implemented.

Although insightful for identifying materials of concern from a hazard perspective, CHA results are difficult to quantify and to aggregate, which makes it difficult to compare technology alternatives. The approach developed in this study provides a framework for data quantification, harmonization, and numerical aggregation so that data gaps and uncertainties can be systematically and transparently evaluated using scenario analysis and decision analysis. Overall, the results indicate that the hazard ranking of LIBs and RFBs are dependent on the current material choices and are sensitive to the variable scenarios applied. The results also change if the mass fraction of materials is considered instead of the number of materials. With the assistance of SMAA, the relative hazard ranking of each battery is more clearly determined, with ZBFB ranking higher in the number-based case and IFB ranking higher in the mass-based case. In contrast, the VRFB ranks lowest regardless of which scenarios are chosen, due to the high hazard and mass fraction of vanadium pentoxide. The hazard ranking for LMO ranks below the other two LIBs due to the use of more BM-1 materials. These findings motivate for the selection of alternative materials with lower hazard, and the quantitative framework allows for an easy, systematic evaluation of these alternatives before being widely implemented.

The combination of the data aggregation strategy and MCDA, as used in this study, provides a systematic but explicit approach for the comparison and selection of safer battery alternatives with overall lower toxicity and hazard. Although demonstrated here to be a viable approach, there are also obstacles encountered such as uncertainties due to the lack of complete and consistent hazard data. Hence, an urgent need to advance the CHA of energy storage system production is to fill the gaps on materials lacking highly reliable and transparent hazard data. Beyond CHA, the MCDA framework could also be expanded to evaluate trade-offs among hazard, cost, battery functionality, carbon footprint, etc., providing a foundation for rigorous assessment of safer material alternatives for these and other energy storage batteries.

CRediT authorship contribution statement

Haoyang He: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Shan Tian:** Investigation, Writing – original draft, **Chris Glaubensklee**: Investigation, **Brian Tarroja:** Writing – review & editing, Supervision, Project administration, Funding acquisition, **Oladele A. Ogunseitan:** Conceptualization, Writing – review & editing, **Scott Samuelsen:** Supervision, Funding acquisition, **Julie M. Schoenung:** Conceptualization, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge funding provided by the California Energy Commission under Agreement #: EPC-16–039. The authors also acknowledge funding from the Lincoln Dynamic Foundation World Institute for Sustainable Development of Materials (WISDOM) at the University of California, Irvine.

Statement of Environmental Implication

Battery energy storage products with a long lifespan such as lithium-ion and redox flow batteries are being installed to support the renewable energy grid. However, the lack of understanding of the inherent toxicity and hazard profiles of the various battery materials will impact the human health and environment in the future. This study closes this gap by presenting a chemical hazard assessment on materials used in battery components and upstream production activities. Moreover, we present a novel, systematic approach to quantitatively evaluate the assessment results, which establishes a critical foundation to inform material selection for future battery design.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jhazmat.2022.129301.

References

Aschberger, K., Campia, I., Pesudo, L.Q., Radovnikovic, A., Reina, V., 2017. Chemical alternatives assessment of different flame retardants – a case study including multi-walled carbon nanotubes as synergist. Environ. Int. 101, 27–45.

Awasthi, A.K., Cheela, V.S., D'Adamo, I., Iacovidou, E., Islam, M.R., Johnson, M., Miller, T.R., Parajuly, K., Parchomenko, A., Radhakrishan, L., Zhao, M., 2021. Zero waste approach towards a sustainable waste management. Res. Environ. Sustain. 3, 100014.

Blum, A.F., Long Jr, R.T., 2016. Fire Hazard Assessment of Lithium Ion Battery Energy Storage Systems. Springer-Verlag, New York.

California Energy Commission, 2021. 2021 SB 100 Joint Agency Report, Achieving 100 Percent Clean Electricity in California: An Initial Assessment. ⟨https://www.energy.ca.gov/publications/2021/2021-sb-100-joint-agency-report-achieving-100-percent-clean-electricity⟩ (Accessed March 23, 2022).

California Senate Bill No. 100, 2017-2018. California Renewables Portfolio Standard Program: emissions of greenhouse gases. 2017–2018.

California Senate Bill No. 32, 2015-2016. California Global Warming Solutions Act of 2006: emissions limit.

Chalamala, B.R., Soundappan, T., Fisher, G.R., Anstey, M.R., Viswanathan, V.V., Perry, M.L., 2014. Redox flow batteries: an engineering perspective. Proc. Ieee. 102 (6), 976–999.

Chen, Y., Kang, Y., Zhao, Y., Wang, L., Liu, J., Li, Y., Liang, Z., He, X., Li, X., Tavajohi, N., Li, B., 2021. A review of lithium-ion battery safety concerns: the issues, strategies, and testing standards. J. Energy Chem. 59, 83–99.

Cinelli, M., Coles, S.R., Kirwan, K., 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. Ecol. Ind. 46, 138–148.

Clean Production Action, 2021a. GreenScreen® for Safer Chemicals. ⟨https://www.greenscreenchemicals.org⟩. (Accessed June 18, 2021).

Clean Production Action, 2021b. GreenScreen CertifiedTM. ⟨https://www.greenscreenchemicals.org/certified⟩ (Accessed October 24, 2021).

Cradle to Cradle Products Innovation Institute, 2021. Cradle to Cradle Certified®. ⟨https://www.c2ccertified.org/get-certified/product-certification⟩ (Accessed October 25, 2021).

da Silva Lima, L., Quartier, M., Buchmayr, A., Sanjuan-Delmás, D., Laget, H., Corbisier, D., Mertens, J., Dewulf, J., 2021. Life cycle assessment of lithium-ion batteries and vanadium redox flow batteries-based renewable energy storage systems. Sustain. Energy Technol. Assess. 46, 101286.

Department of Toxic Substances Control, 2021. Safer Consumer Products. ⟨https://dtsc.ca.gov/scp/⟩ (Accessed July 18, 2021).

Ecoinvent, 2021. ⟨https://www.ecoinvent.org⟩. (Accessed July 20, 2021).

Eisenberg, D.A., Yu, M., Lam, C.W., Ogunseitan, O.A., Schoenung, J.M., 2013. Comparative alternative materials assessment to screen toxicity hazards in the life cycle of CIGS thin film photovoltaics. J. Hazard. Mater. 260, 534–542.

Ellingsen, L.A.W., Majeau-Bettez, G., Singh, B., Srivastava, A.K., Valøen, L.O., Strømman, A.H., 2014. Life cycle assessment of a lithium-ion battery vehicle pack. J. Ind. Ecol. 18 (1), 113–124.

European Chemicals Agency, 2021a. ECHA Registered Substances. ⟨https://echa.europa.eu/information-on-chemicals/registered-substances⟩. (Accessed July 13, 2021).

European Chemicals Agency, 2021b. EU SVHC List. ⟨https://echa.europa.eu/candidate-list-table⟩ (Accessed July 15, 2021).

European Council. 1999. No 1907/2006 of the European Parliament and of the Council of 18 December 2006, concerning the Registration. Eval- uation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and re- pealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Union L. 45:1–849.

Faludi, J., Hoang, T., Gorman, P., Mulvihill, M., 2016. Aiding alternatives assessment with an uncertainty-tolerant hazard scoring method. J. Environ. Manag. 182, 111–125.

Global Electronics Council, 2021. EPEAT Registry. ⟨https://www.epeat.net⟩ (Accessed October 25, 2021).

Harper, G., Sommerville, R., Kendrick, E., Driscoll, L., Slater, P., Stolkin, R., Walton, A., Christensen, P., Heidrich, O., Lambert, S., Abbott, A., 2019. Recycling lithium-ion batteries from electric vehicles. Nature 575 (7781), 75–86.

He, H., Gutierrez, Y., Young, T.M., Schoenung, J.M., 2019a. The role of data source selection in chemical hazard assessment: a case study on organic photocoltaics. J. Hazard. Mater. 365, 227–236.

He, H., Malloy, T.F., Schoenung, J.M., 2019b. Multicriteria decision analysis characterization of chemical hazard assessment data sources. Integr. Environ. Assess. Manag. 15 (6), 895–908.

He, H., Tian, S., Tarroja, B., Ogunseitan, O.A., Samuelsen, S., Schoenung, J.M., 2020. Flow battery production: materials selection and environmental impact. J. Clean. Prod. 269, 121740.

Hiremath, M., Derendorf, K., Vogt, T., 2015. Comparative life cycle assessment of battery storage systems for stationary applications. Environ. Sci. Technol. 49 (8), 4825–4833.

Hu, G., Mian, H.R., Hewage, K., Sadiq, R., 2019. An integrated hazard screening and indexing system for hydraulic fracturing chemical hazard assessment. Process Sat. Environ. Prot. 130, 126–139.

Institute for Occupational Safety and Health of the German Social Accident Insurance, 2021. GESTIS. ⟨https://www.dguv.de/ifa/gestis/gestis-stoffdatenbank/index-2.jsp⟩ (Accessed July 13, 2021).

Jacobs, M.M., Malloy, T.F., Tickner, J.A., Edwards, S., 2016. Alternatives assessment frameworks: research needs for the informed substitution of hazardous chemicals. Environ. Health Perspect. 124 (3), 265–280.

Kalhoff, J., Eshetu, G.G., Bresser, D., Passerini, S., 2015. Safer electrolytes for lithium-ion batteries: state of the art and perspectives. ChemSusChem 8, 2154–2175.

Kang, D.H.P., Chen, M., Ogunseitan, O.A., 2013. Potential environmental and human health impacts of rechargeable lithium batteries in electronic waste. Environ. Sci. Technol. 47, 5495–5503.

Khakzad, N., Reniers, G., 2015. Risk-based design of process plants with regard to domino effects and land use planning. J. Hazard. Mater. 299, 289–297.

Khan, R.A., Anand, A., Wani, M.F., 2018. A holistic framework for environment conscious based product risk modeling and assessment using multi criteria decision making. J. Clean. Prod. 174, 954–965.

Lahdelma, R., Salminen, P., 2001. SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. Oper. Res. 49 (3), 444–454.

Lahdelma, R., Hokkanen, J., Salminen, P., 1998. SMAA-stochastic multiobjective acceptability analysis. Eur. J. Oper. Res. 106 (1), 137–143.

Lam, C.W., Lim, S.R., Ogunseitan, O.A., Shapiro, A.A., Saphores, J.D.M., Brock, A., Schoenung, J.M., 2013. Integrating toxicity reduction strategies for materials and components into product design: a case study on utility meters. Integr. Environ. Assess. Manag. 9 (2), 319–328.

Larcher, D., Tarascon, J.M., 2014. Towards greener and more sustainable batteries for electrical energy storage. Nat. Chem. 7 (1), 19–29.

Leung, P., Li, X., De León, C.P., Berlouis, L., Low, C.J., Walsh, F.C., 2012. Progress in redox flow batteries, remaining challenges and their applications in energy storage. RSC Adv. 2 (27), 10125–10156.

Liang, Y., Su, J., Xi, B., Yu, Y., Ji, D., Sun, Y., Cui, C., Zhu, J., 2017. Life cycle assessment of lithium-ion batteries for greenhouse gas emissions. Resour. Conserve Recycl. 117, 285–293.

Linkov, I., Loney, D., Cormier, S., Satterstrom, F.K., Bridges, T., 2009. Weight-of-evidence evaluation in environmental assessment: review of qualitative and quantitative approaches. Sci. Total Environ. 407 (19), 5199–5205.

Lisbona, D., Snee, T., 2011. A review of hazards associated with primary lithium and lithium-ion batteries. Process Sat. Environ. Prot. 89 (6), 434–442.

Liu, K., Liu, Y., Lin, D., Pei, A., Cui, Y., 2018. Materials for lithium-ion battery safety. Sci. Adv. 4 (6), eaas9820.

Majeau-Bettez, G., Hawkins, T.R., Strømman, A.H., 2011. Life cycle environmental assessment of lithium-ion and nickel metal hydride batteries for plug-in hybrid and battery electric vehicles. Environ. Sci. Technol. 45 (10), 4548–4554.

Malloy, T.F., Zaunbrecher, V.M., Batteate, C.M., Blake, A., Carroll Jr., W.F., Corbett, C.J., Hansen, S.F., Lempert, R.J., Linkov, I., McFadden, R., Moran, K.D., 2017. Advancing alternative analysis: integration of decision science. Environ. Health Perspect. 125 (6), 066001.

National Institute of Technology and Evaluation, 2021. Chemical Management. ⟨https://www.nite.go.jp/chem/english/ghs/ghs_index.html⟩ (Accessed July 12, 2021).

National Library of Medicine, 2021. Hazardous Substances Data Bank, U.S. ⟨https://pubchem.ncbi.nlm.nih.gov/source/11933⟩. (Accessed July 14, 2021).

National Research Council, 2014. A Framework to Guide Selection of Chemical Alternatives. National Academies Press, Washington, DC.

Noack, J., Roznyatovskaya, N., Herr, T., Fischer, P., 2015. The chemistry of redox-flow batteries. Angew. Chem. Int. Ed. 54 (34), 9776–9809.

Notter, D.A., Gauch, M., Widmer, R., Wager, P., Stamp, A., Zah, R., Althaus, H.J., 2010. Contribution of Li-ion batteries to the environmental impact of electric vehicles. Environ. Sci. Technol. 44, 6550–6556.

OECD, 2021. Guidance on Key Considerations for the Identification and Selection of Safer Chemical Alternative. OECD Series on Risk Management, No. 60, Environment, Health, and Safety, Environment Directorate.

Olivetti, E.A., Ceder, G., Gaustad, G.G., Fu, X., 2017. Lithium-ion battery supply chain considerations: analysis of potential bottlenecks in critical metals. Joule 1 (2), 229–243.

Prat, D., Wells, A., Hayler, J., Sneddon, H., McElroy, C.R., Abou-Shehada, S., Dunn, P.J., 2016. CHEM21 selection guide of classical- and less classical-solvent. Green. Chem. 18, 288.

Schmidt, T.S., Beuse, M., Zhang, X., Steffen, B., Schneider, S.F., Pena-Bello, A., Bauer, C., Parra, D., 2019b. Additional emissions and cost from storing electricity in stationary battery systems. Environ. Sci. Technol. 53 (7), 3379–3390.

Schmidt, O., Melchior, S., Hawkes, A., Staffell, I., 2019a. Projecting the future levelized cost of electricity storage technologies. Joule 3 (1), 81–100.

Song, J., Yan, W., Cao, H., Song, Q., Ding, H., Lv, Z., Zhang, Y., Sun, Z., 2019. Material flow analysis on critical raw materials of lithium-ion batteries in China. J. Clean. Prod. 215, 570–581.

Tarroja, B., Zhang, L., Wifvat, V., Shaffer, B., Samuelsen, S., 2016. Assessing the stationary energy storage equivalency of vehicle-to-grid charging battery electric vehicles. Energy 106, 673–690.

Tenaw, D., Hawitibo, A.L., 2021. Carbon decoupling and economic growth in Africa: evidence from production and consumption-based carbon emissions. Resour. Environ. Sustain 6, 100040.

Tervonen, T., 2014. JSMAA: open source software for SMAA computations. Int. J. Syst. Sci. 45 (1), 69–81.

Swedish Confederation of Professional Employees, 2021. TCO Certified. ⟨https://tcocertified.com⟩ (Accessed October 25, 2021).

The Endocrine Disruption Exchange, 2021. ⟨https://endocrinedisruption.org⟩ (Accessed July 16, 2021).

Tian, S., He, H., Kendall, A., Davis, S.J., Ogunseitan, O.A., Schoenung, J.M., Samuelsen, S., Tarroja, B., 2021. Environmental benefit-detriment thresholds for flow battery energy storage systems: a case study in California. Appl. Energy 300, 117354.

U.S. Environmental Protection Agency, 2021c. EPI SUITETM – Estimation Program Interface. ⟨https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface⟩ (Accessed July 15, 2021).

United Nations, 2019. Globally harmonized system of classification and labelling of chemicals (GHS). Eighth revised edition. New York and Geneva. ⟨https://unece.org/ghs-rev8–2019⟩ (Accessed June 8, 2021).

United States Environmental Protection Agency, 2021a. Chemical under the Toxic Substances Control Act. ⟨https://www.epa.gov/chemicals-under-tsca⟩. (Accessed July 14, 2021).

United States Environmental Protection Agency, 2021b. Design for the Environment Alternative Assessments. ⟨https://www.epa.gov/saferchoice/design-environment-alternatives-assessments⟩. (Accessed July 16, 2021).

Weber, S., Peters, J.F., Baumann, M., Weil, M., 2018. Life cycle assessment of a vanadium redox flow battery. Environ. Sci. Technol. 52 (18), 10864–10873.

Wood III, D.L., Li, J., Daniel, C., 2015. Prospects for reducing the processing cost of lithium ion batteries. J. Power Source 275, 234–242.

Zanghelini, G.M., Cherubini, E., Soares, S.R., 2018. How multi-criteria decision analysis (MCDA) is aiding life cycle assessment (LCA) in results interpretation. J. Clean. Prod. 172, 609–622.

Zheng, Z., Peters, G.M., Arp, H.P.H., Andersson, P.L., 2019. Combining in silico tools with multicriteria analysis for alternatives assessment of hazardous chemicals: a case study of decabromodiphenyl ether alternatives. Environ. Sci. Technol. 53 (11), 6341–6351.

Zubi, G., Dufo-López, R., Carvalho, M., Pasaoglu, G., 2018. The lithium-ion battery: State of the art and future perspectives. Renew. Sustain. Energy Rev. 89, 292–308.