

UCLA

UCLA Previously Published Works

Title

Has Toxicity Testing Moved into the 21st Century? A Survey and Analysis of Perceptions in the Field of Toxicology

Permalink

<https://escholarship.org/uc/item/7m16b52q>

Journal

Environmental Health Perspectives, 125(8)

ISSN

0091-6765

Authors

Zaunbrecher, Virginia

Beryt, Elizabeth

Parodi, Daniela

et al.

Publication Date

2017-08-16

DOI

10.1289/ehp1435

Peer reviewed

Has Toxicity Testing Moved into the 21st Century? A Survey and Analysis of Perceptions in the Field of Toxicology

Virginia Zaunbrecher,^{1,2} Elizabeth Beryt,³ Daniela Parodi,⁴ Donatello Telesca,⁵ Joseph Doherty,² Timothy Malloy,^{1,2,6} and Patrick Allard^{1,4,6}

¹Sustainable Technology and Policy Program, University of California, Los Angeles, Los Angeles, California, USA

²School of Law, University of California, Los Angeles, Los Angeles, California, USA

³Luskin School of Public Affairs, University of California, Los Angeles, Los Angeles, California, USA

⁴Institute for Society and Genetics, University of California, Los Angeles, Los Angeles, California, USA

⁵Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California, USA

⁶Department of Environmental Health Sciences, UCLA Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California, USA

BACKGROUND: Ten years ago, leaders in the field of toxicology called for a transformation of the discipline and a shift from primarily relying on traditional animal testing to incorporating advances in biotechnology and predictive methodologies into alternative testing strategies (ATS). Governmental agencies and academic and industry partners initiated programs to support such a transformation, but a decade later, the outcomes of these efforts are not well understood.

OBJECTIVES: We aimed to assess the use of ATS and the perceived barriers and drivers to their adoption by toxicologists and by others working in, or closely linked with, the field of toxicology.

METHODS: We surveyed 1,381 toxicologists and experts in associated fields regarding the viability and use of ATS and the perceived barriers and drivers of ATS for a range of applications. We performed ranking, hierarchical clustering, and correlation analyses of the survey data.

RESULTS: Many respondents indicated that they were already using ATS, or believed that ATS were already viable approaches, for toxicological assessment of one or more end points in their primary area of interest or concern (26–86%, depending on the specific ATS/application pair). However, the proportions of respondents reporting use of ATS in the previous 12 mo were smaller (4.5–41%). Concern about regulatory acceptance was the most commonly cited factor inhibiting the adoption of ATS, and a variety of technical concerns were also cited as significant barriers to ATS viability. The factors most often cited as playing a significant role (currently or in the future) in driving the adoption of ATS were the need for expedited toxicology information, the need for reduced toxicity testing costs, demand by regulatory agencies, and ethical or moral concerns.

CONCLUSIONS: Our findings indicate that the transformation of the field of toxicology is partly implemented, but significant barriers to acceptance and adoption remain. <https://doi.org/10.1289/EHP1435>

Introduction

Toxicological assessment is central to ensuring safety and protection from potentially dangerous substances. Yet, there is limited toxicological information available for thousands of chemicals currently present in consumer products, industrial processes, food, and the environment (Gomez 2013; Neltner et al. 2013; Stephenson 2009; Tice et al. 2013). Conventional testing relies mainly on whole-animal studies—dosing animals, usually mammals, with a compound and comparing health end points with a control group (National Research Council 2007). However, this approach has been criticized for being slow and expensive, for failing to evaluate all critical end points and life stages (Tice et al. 2013), for inaccurately predicting human health impacts because of species-specific responses to compounds (Leist and Hartung 2013), and for heavy animal use (Tice et al. 2013). These limitations have led the National Academy of Sciences (NAS), as well as governmental agencies, scientists, and stakeholders across the globe, to call for a rethinking of chemical safety evaluation (Judson et al. 2009; National Research Council 2007).

The demand for accurate and efficient toxicology testing, and the advent of new techniques in molecular and computational biology, spurred the development of alternative testing strategies (ATS). These are primarily *in vitro* and *in silico* methodologies that evaluate changes in biological processes at the molecular, cellular, and tissue levels. For example, high-throughput screening (HTS) approaches use advanced robotics and automation to test hundreds to thousands of samples at a time (National Research Council 2007). Read-across and quantitative structure–activity relationship model (QSAR) approaches can be used to estimate the toxicity of data-poor chemicals based on data from similar chemicals (National Research Council 2007). Such a combination of approaches has recently been evaluated and was deemed acceptable by the Endocrine Disruptor Screening Program of the U.S. Environmental Protection Agency (EPA) as a valuable alternative to several Tier 1 assays for chemicals with estrogen activity (Browne et al. 2015). However, the initial vision put forth in the NAS report was also met with some reticence and pessimism with regard to the challenges that would need to be overcome for its implementation (Andersen and Krewski 2010). Furthermore, a decade after the NAS acknowledged the transformational role ATS could play in risk management, little is known about how widespread the use of ATS is and what could be done to accelerate its acceptance.

We addressed this gap by conducting a survey of scientists and nonscientists whose training and/or occupation reside in, or are closely linked with, the field of toxicology. The survey aimed to answer three main questions: *a*) What is the current degree of acceptance of ATS? *b*) What are the barriers to the widespread adoption of ATS? *c*) What factors promote their adoption? We developed the survey questions by drawing upon a literature review and upon 19 semistructured interviews of experts from industry, government, nongovernmental organizations, and academia in the United States and Europe. The online questionnaire,

Address correspondence to P. Allard, Institute for Society and Genetics, 3360 Life Science Building, University of California, Los Angeles, Los Angeles, CA 90095 USA. Telephone: (310) 825-5257. Email: pallard@ucla.edu

Supplemental Material is available online (<https://doi.org/10.1289/EHP1435>).

The authors declare they have no actual or potential competing financial interests.

Received 29 November 2016; Revised 15 June 2017; Accepted 17 June 2017; Published 30 August 2017.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

disseminated through toxicological professional societies and trade groups, was completed by 1,381 respondents whose responses provided information on their perceptions of the use and utility of ATS for a range of applications and settings.

To our knowledge, this is the first comprehensive summary of toxicologists' opinions about the state and utility of alternative toxicity testing since the NAS issued a 2007 report calling for a paradigm shift in toxicology (National Research Council 2007). We report survey findings regarding the nature and scope of ATS use within private businesses and public agencies, and respondents' perceptions of the scientific and sociolegal barriers to and drivers of broader adoption. We conclude with a discussion of the likely impact of recent reforms of the federal Toxic Substances Control Act on the adoption of ATS for regulatory purposes.

Methods

Survey Design

Survey design began with semistructured interviews conducted by phone with 19 toxicology experts between November 2013 and March 2014. Participants included toxicologists and other experts working for state and national governments, various sectors of industry, and nonprofit groups in the United States and Europe. The interviews were conducted using a prepared set of interview topics and themes, but leaving open the opportunity to explore additional topics or information that arose during the interview. Topics included the respondents' overall knowledge of alternative testing approaches; their use of these alternative testing approaches; their perception of the viability, advantages, and drawbacks of these alternative testing approaches; their views regarding the use of data generated by alternative testing approaches into risk assessment, management, and the regulatory context; and their perception of the future of the field pertaining to these alternative testing approaches.

The digitally recorded interviews were summarized by topic area in a Microsoft Excel spreadsheet, and the results were used to generate a draft version of the survey, along with information from the literature and the experience of the project team. The draft survey was piloted with three experts in toxicology testing and regulation drawn from business, government, and academia. Each pilot respondent completed the entire survey, providing contemporaneous comments regarding the survey design, including comments regarding structure, terminology, clarity, and scope. The comments were recorded and reviewed by the project team, and revisions to the survey were made where necessary to address concerns reflected in the comments. A PDF version of the final online survey can be accessed in the Supplemental Material ("Survey and Answers").

The survey primarily focused upon six types of ATS identified in the literature (Collins et al. 2008; Leung et al. 2008; National Research Council 2007; Yang et al. 2009), which were described in the survey as follows: mechanistically based *in vitro* cell or biochemical assays, high-throughput *in vitro* cell or biochemical assays, mechanistically based *in vivo* small-animal assays (e.g., zebrafish or *C. elegans*), high-throughput *in vivo* small-animal assays (e.g., zebrafish or *C. elegans*), QSARs, and biomarkers. Because the survey was directed at toxicologists and experts in related fields, no definitions of the ATS types were included in the survey. For readers who may be unfamiliar with one or more ATS types, descriptions of the ATS types derived from the relevant literature are provided in the Supplemental Material (see Table S1). Respondents were asked to assess the current and future viability of each ATS type for seven specified applications for one or more end points in their respective

primary areas of interest or concern: screening/prioritization for further testing; screening/prioritization for other actions (e.g., risk assessment, risk management); setting doses for *in vivo* testing; weight of evidence in quantitative risk assessment (scoping to determine the most sensitive end points); qualitative risk assessment (e.g., control banding); quantitative risk assessment [identifying the no observed adverse effects level (NOAEL) or other levels]; comparative assessment of alternative chemicals/products/processes (alternatives analysis). Again, because the survey was directed at toxicologists and experts in related fields, no definitions of the potential applications were included in the survey (see Table S2 for descriptions of the applications). We also queried respondents regarding specified sets of potential barriers to and drivers of adoption of each of the six ATS types. Additionally, the survey collected information regarding the respondent's gender, age, country of work, employer type, industry sector, current position, and education (Table 1), which

Table 1. Demographic information of survey respondents.

Characteristic	n (%)
Gender	1,310 ^a
Female	481 (36.7)
Male	826 (63.0)
Other	3 (0.2)
Did not answer	71
Year of most recent degree	1,302
2010–present	166 (12.7)
2000–2009	329 (25.3)
1990–1999	350 (26.7)
1980–1989	291 (22.4)
1970–1979	129 (9.9)
Before 1970	37 (2.8)
Did not answer	79
Geographic region	1,324
Europe	201 (15.2)
North America	991 (74.8)
Other	132 (10.0)
Did not answer	57
Degrees held ^b	1,319
Undergraduate	557 (42.2)
Masters	459 (34.8)
Doctorate	1,037 (78.6)
Medical	59 (4.5)
Law	5 (0.4)
Other	71 (5.4)
Did not answer	62
Employer	1,381
Academia/research institute	403 (29.2)
Large business	318 (23.0)
Small/medium business	89 (6.4)
National government	180 (13.0)
State/local government	53 (3.8)
Other	338 (24.5)
Did not answer	0
Sector ^c	1,328
Pharmaceuticals only	271 (20.4)
Other	1,057 (79.6)
Did not answer	53

^aIndicates the total number of responses to each demographic question.

^bRespondents were asked, "In which of the following countries do you primarily work?" Respondents could select more than one country. Respondents who only selected North American countries were classified as "North America." Respondents who selected any European country were classified as "Europe." Respondents who selected neither North American nor European countries but did select other countries were classified as "Other."

^cRespondents were asked to "identify what degree(s) you hold." More than one degree could be indicated. Percentages indicate the number of people who selected each degree over the total number of people who responded to the question.

^dRespondents were asked, "To which industry sectors, if any, does your work relate?" Respondents could select more than one sector. Numbers and percentages reflect respondents who only selected "pharmaceuticals."

allowed us to compare the sample frame with the Society of Toxicology's (SoT's) demographics (see Table S3).

We conducted the survey between 26 September 2014 and 14 November 2014, with 95% of the responses received between the start date and 30 October 2014. The responses ($n = 1,381$) were collected using the online survey application Qualtrics (<https://www.qualtrics.com/>). The sampling frame consisted of members of professional societies most closely associated with toxicology: namely, SoT, the Society of Environmental Toxicology and Chemistry (SETAC), the Center for Alternatives to Animal Testing (CAAT), the American Society for Cellular and Computational Toxicology (ASCCT), and AltTox. Currently active toxicologists worldwide are likely to belong to at least one of these organizations. SoT e-mailed one recruitment and one follow-up e-mail to each of its members on our behalf, with an approximate response rate of 17% ($n = 936$). An e-mail announcement of the survey along with a single URL was distributed through both SETAC and CAAT member communications, resulting in 462 additional responses. Separate links were also sent to mailing lists of toxicologists at the U.S. EPA ($n = 7$), ASCCT ($n = 18$), and AltTox ($n = 11$), resulting in 36 additional responses to the survey. The response rates for SETAC, CAAT, U.S. EPA, ASCCT, and AltTox cannot be calculated given the nature of the recruitment process and the fact that the membership lists of these associations and organizations overlap. Participants were given the option of providing their name and e-mail address at the end of the survey if they were willing to be contacted for follow-up questions; otherwise, their responses were anonymous.

The survey was designed so that individual items may not be relevant to all respondents. Consequently, the number of responses to individual questions varied. All participants' personal information was confidential, and survey administrators and analysts were blinded to the identity of the respondents as described in the Institutional Review Board (IRB) exemption granted by University of California, Los Angeles (UCLA) (IRB No. 13-001339). The summary of the responses without identifiers can be found in the Supplemental Material (see "Survey and Responses").

Statistical and Clustering Analyses

We examined respondents' answers for four categories of questions: familiarity with ATS (question 12), current use and perceived viability of each of six ATS approaches (questions 16–21), perceived barriers to viability or adoption (questions 23 and 24), and current or future drivers of adoption by the respondent or by their organization (question 25). We used a two-step analysis to identify meaningful differences among groups of respondents for each of those categories. In the first step, we clustered individuals using hierarchical clustering, a statistical method that aims to recursively partition a set of multivariate measurements on the basis of a dissimilarity measure. The technique mines the data to identify nested groups of respondents who provided similar answers, resulting in an easily interpretable dendrogram configuration of the original data. To accomplish this aim, we generated groups of subjects using the agglomerative version of Ward's complete linkage minimum variance algorithm with Manhattan distance matrices, and our analysis was implemented using the R function "hclust" (R Project for Statistical Computing). In all applications, the "other" category was excluded owing to a high prevalence of nonresponse. In all other categories, nonresponse was only excluded in the computation of pairwise distances. For a conservative interpretation of the dissimilarities among the respondents, we broke the respondents into two major clusters for each of the categories of questions, which separated the

respondents into two easily interpretable categories. For example, cluster analysis of the viability of mechanistically based *in vitro* assays generated two clusters of respondents, those perceiving high viability and those perceiving low viability. A detailed example of how hierarchical clustering identifies latent classes of survey respondents is illustrated for overall viability in Figure S1.

In the second analytical stage, we associated cluster membership with several of the respondents' demographic characteristics using a logistic regression model: year of terminal degree (modeled as a continuous variable), employer type (national government, state government, large business, small/medium business, academia, and other), geographic association (Europe, North America, and other), gender (M, F), and sector (selected Pharma as only choice vs. did not select it). In all of these analyses, we used males, North American, employment in a National Government as the baseline/reference categories. The logistic regression model was assessed for adequacy, comparing its classification results with random forests (RF). Specifically, we compared cross-validation estimates (10CV) of classification accuracy for both RF and logistic regression. For all models, logistic regression models are at most 5% less accurate than the RF estimate (Liaw and Wiener 2002). The data were processed for analysis in three steps: cleaning up of the data, analysis, and report generation. In our assessment of statistical significance, we controlled type I errors (α) at or below 5%. Links to the complete data set, data cleaning routing, the data analysis routines, and the report-generating codes are available in the Supplemental Material (see "Data and Code").

Finally, the analysis of perception of viability by familiarity with ATS (see Table S4) was performed by tabulating the percentage of survey respondents, classified by hierarchical clustering as perceiving specific ATS strategies as being currently viable, by degree of familiarity with specific ATS strategies. *p*-Values are based on Chi-squared tests of independence. Controlling type I errors at 5%, a *p*-value <0.05 identified a significant association between perceived viability and familiarity categories. The analysis of respondents' characteristics with regard to familiarity (see Table S5) was performed by categorizing familiarity into four classes: namely, unfamiliar (unfamiliar with all listed technologies), somewhat familiar (somewhat familiar with fewer than half of the listed technologies), familiar (somewhat or very familiar with more than half, but not all, of the listed technologies), and very familiar (somewhat or very familiar with all listed technologies). The data were then described as marginal associations between familiarity and gender, degree year, employer, region, and pharma. The reported *p*-values refer to Chi-squared tests of independence for categorical predictors and to analysis of variance (ANOVA) *F*-tests for continuous predictors (degree year).

Results

The survey respondents varied in age, professional background, and geographic location. The majority of respondents were members of SoT (68% of total respondents), our primary survey dissemination channel, and from the United States (73%). Nearly a quarter of the respondents (23%) were from European Union countries. Further demographics characterization is found in Table 1. To ascertain the representativeness of the respondent population, we compared the demographics of the SoT membership at the time of the survey with that of our respondent population. The sample demographics aligned closely with the SoT demographics with only three minor differences: the study sample population was skewed toward people who obtained their last degree more recently than the full SoT membership; government respondents were over-represented (17% from the sample vs. 13% from SoT);

and industry was under-represented (30% from the sample vs. 36% from SoT) (see Table S3).

Use and Viability of Alternative Testing Strategies

Many respondents already perceive ATS as viable for a variety of applications; that is, respondents either use ATS for a relevant application or believe that it can be used currently for that application. However, perception of viability varied considerably by the intended use of the ATS (Table 2). For example, ATS for screening or prioritizing chemicals for further testing had the widest acceptance; depending upon the particular ATS approach, 70.1% to 86.4% of respondents viewed ATS to be currently acceptable. Acceptance of using ATS for establishing levels at which toxic effects may occur (i.e., NOAELs) in quantitative risk assessment (QRA) varied by ATS approach. Acceptance rates were highest for QSARs (42.8%) and biomarkers (53.5%) and lowest for HTS *in vitro* technologies (25.5%). Similarly, a majority of respondents (69.2–53.2% depending upon the ATS approach) considered ATS to be viable for the comparative assessment of alternative chemicals.

The apparent support for current viability was tempered by other survey results. Particularly with respect to its application to QRA, notable numbers of respondents perceived most ATS

approaches as neither currently viable nor viable in the foreseeable future (Table 3). For example, 25.7% of respondents viewed QSARs as unfeasible for QRA; more than a third held the same view of *in vitro* HTS. The biomarkers category was the exception, with just 14.7% of respondents being skeptical. Moreover, there was a gap between acceptance and actual use. Although many respondents judged these technologies to be currently viable for a variety of applications, fewer respondents reported using the technologies for those applications. In terms of current use, screening for further testing using mechanistic *in vitro* approaches was most common (41.1%), whereas employing HTS *in vivo* for qualitative risk assessment was least common (4.5%) (Table 2). There is some indication that part of this difference between perceived viability and actual use may be related to regulatory constraints; as we discuss below, concern about regulatory acceptance is one of the leading barriers to adoption. However, a variety of other factors may also affect the respondent's use of a given method, including whether the particular ATS/application is relevant to end points in their primary area of interest or concern and the scope of the respondent's job responsibilities.

Hierarchical clustering generated two groups of respondents based on survey questions about the perceived viability of ATS methods in a broad range of technologies and applications (Q14

Table 2. Use and viability: numbers (%) of participants who indicated that an alternative testing strategy technology is a viable approach for a toxicological assessment (application) of one or more end points in their primary area of interest or concern.

Application	ATS Technology ^a n (%)					
	Mechanistic <i>in vitro</i>	HTS <i>in vitro</i>	Mechanistic <i>in vivo</i>	HTS <i>in vivo</i>	QSARs	Biomarkers
Screening/prioritization for further testing	884 ^b	851	790	738	942	845
Current user ^c	363 (41.1)	242 (28.4)	128 (16.2)	90 (12.2)	334 (35.5)	298 (35.3)
Currently viable ^d	379 (42.9)	452 (53.1)	458 (58.0)	427 (57.9)	480 (51.0)	399 (47.2)
Currently acceptable ^e	712 (83.9)	694 (81.6)	586 (74.2)	517 (70.1)	814 (86.4)	697 (82.5)
Screening/prioritization for other actions	859	827	773	720	911	846
Current user	273 (31.8)	156 (18.9)	98 (12.7)	70 (9.7)	309 (33.9)	264 (31.2)
Currently viable	343 (39.9)	404 (48.9)	400 (51.7)	356 (49.4)	389 (42.7)	387 (45.7)
Currently acceptable	616 (71.7)	560 (67.7)	498 (64.4)	426 (59.2)	698 (76.6)	651 (77.0)
Qualitative risk assessment	721	694	656	601	749	712
Current user	137 (19.0)	59 (8.5)	52 (7.9)	27 (4.5)	152 (20.3)	154 (21.6)
Currently viable	266 (36.9)	255 (36.7)	285 (43.4)	232 (38.6)	317 (42.3)	323 (45.4)
Currently acceptable	403 (55.9)	314 (45.2)	337 (51.4)	259 (43.1)	469 (62.6)	477 (67.0)
Setting doses for <i>in vivo</i> testing	824	764	729	693	842	781
Current user	145 (17.6)	70 (9.2)	76 (10.4)	49 (7.1)	137 (16.3)	195 (25.0)
Currently viable	265 (32.2)	212 (27.7)	254 (34.8)	214 (30.9)	306 (36.3)	304 (38.9)
Currently acceptable	410 (49.8)	282 (36.9)	330 (45.3)	263 (38.0)	443 (52.6)	499 (63.9)
Weight of evidence in quantitative risk assessment	828	783	745	689	874	799
Current user	187 (22.6)	80 (10.2)	78 (10.5)	41 (6.0)	229 (26.2)	209 (26.2)
Currently viable	300 (36.2)	289 (36.9)	310 (41.6)	242 (35.1)	336 (38.4)	341 (42.7)
Currently acceptable	487 (58.8)	369 (47.1)	388 (52.1)	283 (41.1)	565 (64.6)	550 (68.8)
Setting NOAEL or other levels in quantitative risk assessment	799	758	743	684	873	774
Current user	100 (12.5)	44 (5.8)	66 (8.9)	39 (5.7)	176 (20.2)	164 (21.2)
Currently viable	162 (20.3)	149 (19.7)	179 (24.1)	145 (21.2)	198 (22.7)	250 (32.3)
Currently acceptable	262 (32.8)	193 (25.5)	245 (33.0)	184 (26.9)	374 (42.8)	414 (53.5)
Comparative assessment of alternatives	798	764	725	671	857	762
Current user	191 (23.9)	97 (12.7)	76 (10.5)	54 (8.0)	205 (23.9)	171 (22.4)
Currently viable	335 (42.0)	346 (45.3)	344 (47.4)	303 (45.2)	388 (45.3)	351 (46.1)
Currently acceptable	526 (65.9)	443 (58.0)	420 (57.9)	357 (53.2)	593 (69.3)	522 (68.5)

Note: ATS, alternative testing strategies; HTS, high-throughput screening; NOAEL, no observable adverse effects level; QSAR, quantitative structure–activity relationship. For each technology, respondents were asked, “To what extent do you believe that the use of [the technology] is a viable approach for the following aspects of toxicological assessment for one or more end points in your primary area of interest or concern?” For each application, respondents were given the following choices: “I have used it for this purpose in the last 12 months,” “Is a viable use, but I have not used it for this purpose in the last 12 months,” “Not currently viable, but may be viable within 1–5 years,” “Not currently viable, but may be viable within 5–10 years,” “Not a viable use now or in the foreseeable future,” and “Do not know/not sure.” Percentages for each technology/application pair are based on the number of respondents who answered the question, excluding “Do not know/not sure” responses.

^aSurvey questions used the following terms to describe the six ATS technologies: Mechanistic *in vitro*, mechanistically based *in vitro* cell or biochemical assays; HTS *in vitro*, high-throughput screening *in vitro* cell or biochemical assays; mechanistic *in vivo*, mechanistically-based *in vivo* cell or small-animal assays (e.g., zebrafish or *C. elegans*); HTS *in vivo*, high-throughput screening *in vivo* small-animal assays (e.g., zebrafish or *C. elegans*); QSARs, quantitative structure activity relationship models; biomarkers, biomarkers].

^bIndicates the total number of responses for each technology/application combination.

^cCurrent users indicated, “I have used it for this purpose in the last 12 months.”

^d“Currently viable” means respondents indicated, “Is a viable use, but I have not used it for this purpose in the last 12 months.”

^e“Currently acceptable is the sum of “Current user” and “Currently viable” for each technology/application combination.

Table 3. Nonviable alternative testing strategies: numbers (%) of respondents who indicated that an ATS technology is not a viable approach for a toxicological assessment (application) of one or more end points in their primary area of interest or concern.

Application	Mechanistic <i>in vitro</i>	HTS <i>in vitro</i>	Mechanistic <i>in vivo</i>	HTS <i>in vivo</i>	QSARs	Biomarkers
Screening/prioritization for further testing	26/884 (2.9)	38/851 (4.5)	56/790 (7.1)	70/738 (9.5)	26/942 (2.8)	33/845 (3.9)
Screening/prioritization for other actions	51/859 (5.9)	77/827 (9.3)	78/773 (10.1)	90/720 (12.5)	43/911 (4.7)	42/846 (5.0)
Qualitative risk assessment	83/721 (11.5)	114/694 (16.4)	113/656 (17.2)	120/601 (20.0)	60/749 (8.0)	61/712 (8.6)
Setting doses for <i>in vivo</i> testing	134/824 (16.3)	191/764 (25.0)	167/729 (22.9)	187/693 (27.0)	146/842 (17.3)	78/781 (10.0)
Weight of evidence in quantitative risk assessment	94/828 (11.4)	127/783 (16.2)	127/745 (17.0)	156/689 (22.6)	75/874 (8.6)	65/799 (8.1)
Quantitative risk assessment (identifying NOAEL or other levels)	213/799 (26.7)	264/758 (34.8)	230/743 (31.0)	243/684 (35.5)	224/873 (25.7)	114/774 (14.7)
Comparative assessment of alternatives	59/798 (7.4)	85/764 (11.1)	101/725 (13.9)	106/671 (15.8)	60/857 (7.0)	54/762% (7.1)

Note: ATS, alternative testing strategies; HTS, high-throughput screening; NOAEL, no observable adverse effects level; QSAR, quantitative structure–activity relationship. Survey questions were the same as those in Table 2. Percentages for each technology/application pair are based on the number of respondents who selected “Not a viable use now or in the foreseeable future” divided by the total number of respondents for that technology/use combination excluding “do not know/not sure” responses.

to Q19) (see Figure S1). A post hoc examination of the clustering results identified one group of respondents perceiving most applications of ATS as being highly viable in the short term (90% of respondents see most applications as potentially viable currently or within the next 5 y), and the comparison group with a more critical attitude towards the short-term viability (50% of respondents identifying only between 45% and 84% of ATS applications as viable currently or within the next 5 y). Logistic regression analyses identified that degree year was modestly positively correlated with perceived viability of technologies [OR for a 10-y increase = 1.03 (95% CI: 1.02, 1.05); $p = 0.001$] (Table 4). There was no statistical difference between the views of Europeans and North Americans. Compared with respondents working in all other sectors combined, those working in the pharmaceutical industry were less likely to view ATS as viable, although the difference was not significant [OR = 0.92 (95% CI: 0.65, 1.30); $p = 0.642$].

Finally, we investigated whether familiarity was a potential factor behind the perception of viability. The more familiar a respondent was with a given technology, the more likely the respondent was to perceive it as currently viable for at least one use (see Table S4). This factor appears to be particularly important for recent graduates, for people outside of the United States, and for employees of small/medium businesses who are all on balance less familiar with alternative testing approaches (see Table S5).

Barriers to Adoption of Alternative Testing Strategies

The survey included two questions related to barriers to ATS. The first question asked participants to indicate, for each of the alternative approaches that they believed was not currently viable, the

factors they saw as significant barriers to viability. Possible responses to this question included 11 scientific or technical factors that might be perceived as barriers (Figure 1A). The second question asked participants to indicate which of 13 possible social/legal/institutional barriers they thought played a significant role in inhibiting adoption of the six ATS by the respondent or by others in their organization (Figure 1B). By total number of responses across the six ATS, scientific/technical barriers accounted for 4 of the 5 most frequent responses to the two questions combined. These technical barriers included concerns with regard to the interpretation and extrapolation of the data, failure to capture the integrated whole-animal system, the difficulty in developing dose–response relationships, and concerns with the accuracy in terms of false positives and false negatives (Figure 1A). The rank order of the barriers was generally consistent across the approaches, with two exceptions. The failure to capture the integrated whole-animal system was the leading barrier for three approaches: mechanistically based *in vitro* assays, HTS *in vitro* assays, and QSARs. This barrier dropped to eighth for biomarkers and to ninth for mechanistically based *in vivo* small animal assays and HTS *in vivo* small animal assays. In addition, difficulties in management/synthesis of large amounts of data were a leading barrier for high-throughput testing approaches, but not for the other testing approaches. Notably, the two least commonly indicated scientific/technical barriers were “existing testing approaches are adequate” and “not scientifically sound.”

In the next question, we inquired about social/legal/institutional barriers that may inhibit ATS adoption by the respondents or by their organization. By a wide margin, respondents indicated “concern about regulatory acceptance” as the most common factor inhibiting the adoption of ATS (2,192 positive responses), followed by “lack of scientific validation” (1,571 responses) (Figure 1B). Concerns regarding the current validity and the validation process for ATS methods were the next-most-prominent perceived nontechnical barriers, including the complexity and pace of validation as well as the availability of validated ATS methods. Validation is the process by which “the reliability and relevance of a new method is established for a specific purpose” (OECD 2005). Another significant social/legal/institutional barrier is general resistance to change, which is addressed in more detail in the discussion below.

Hierarchical clustering generated two groups of respondents based on survey questions about factors that they perceived as significant barriers to the viability of one or more ATS (Q21) and factors that they believed play a significant role in inhibiting the adoption of ATS (Q22), with one group identifying more barriers (50% of respondents identifying between 39 and 67 barriers) and the comparison group identifying fewer barriers (50% of respondents identifying between 12 and 27 barriers). In our analysis, we

Table 4. Logistic regression: Overall viability.

Characteristic	OR (95% CI)	Natural log		
		OR	SE	p -Value
Gender (F) ^a	1.07 (0.82, 1.39)	0.07	0.14	0.614
Degree year/10 y	1.03 (1.02, 1.05)	0.03	0.01	0.001
Region (Europe)	1.11 (0.78, 1.57)	0.10	0.18	0.564
Region (other)	1.01 (0.66, 1.53)	0.01	0.22	0.975
Employer (small/medium business)	1.58 (0.85, 2.94)	0.46	0.32	0.152
Employer (large business)	1.35 (0.84, 2.18)	0.30	0.24	0.215
Employer (state government)	1.91 (0.90, 4.03)	0.65	0.38	0.091
Employer (academia)	1.66 (1.04, 2.65)	0.51	0.24	0.033
Employer (other)	1.74 (1.11, 2.74)	0.56	0.23	0.016
Sector (Pharma)	0.92 (0.65, 1.30)	−0.08	0.18	0.642

Note: ATS perceived as viable explained by respondents’ characteristics compared with a reference group of U.S. males working for a national government organization and not associated with the pharmaceutical industry. ATS, alternative testing strategies; CI, confidence interval; OR, odds ratio; SE, standard error.

^aThree subjects classified as “Other” are not included in this analysis.

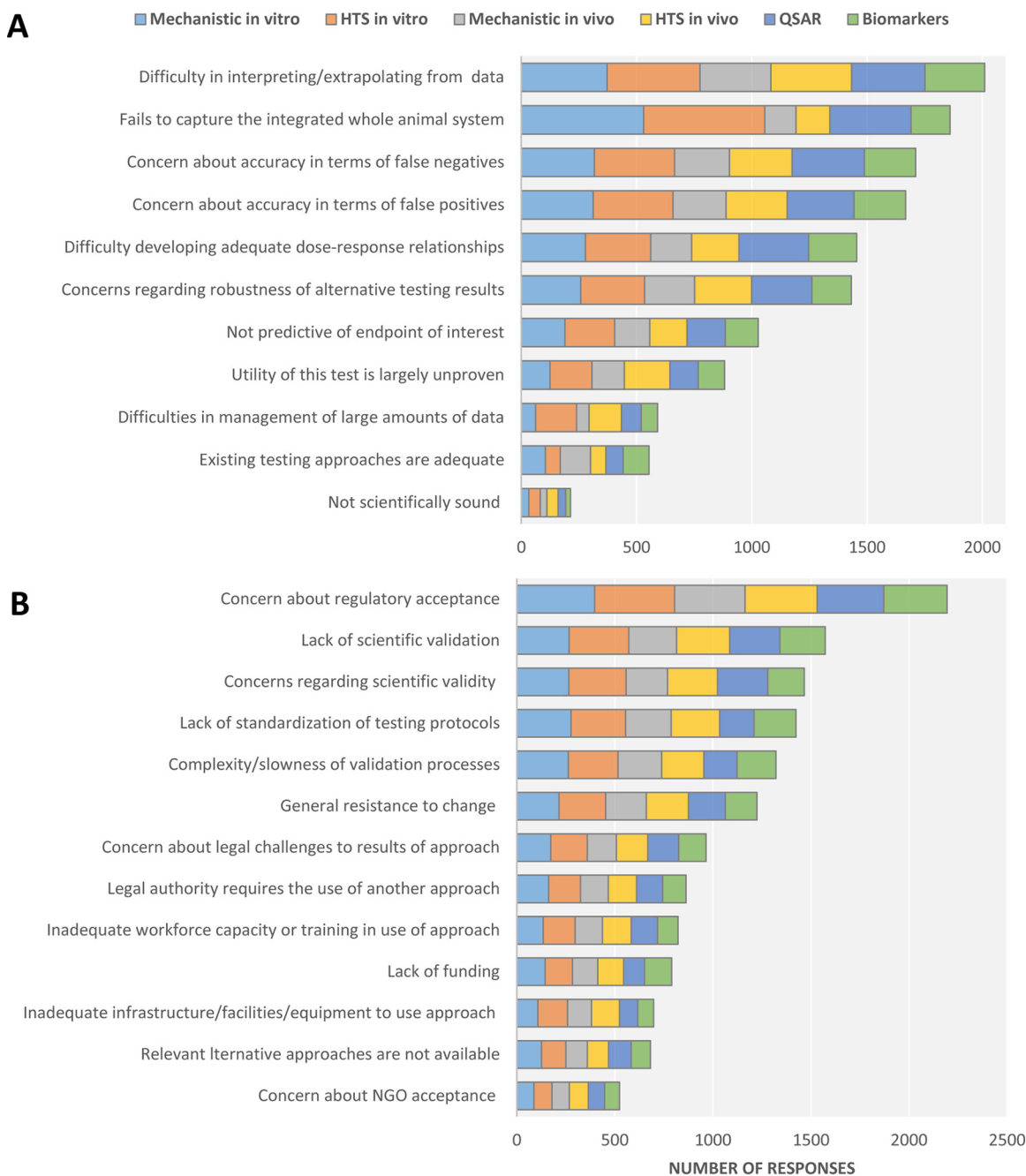


Figure 1. Social and institutional barriers. (A) Respondents were asked to identify, for each alternative approach they believe is not currently viable, which factors they see as significant barriers to viability (Question 21). (B) Additionally, the respondents were asked to identify which factors, if any, they think play a significant role in inhibiting the adoption of the listed alternative approaches by them or by others in their organization (Question 22). For both panels, the y-axis represents the total number of responses collected and is further divided by category of alternative technology. Respondents were able to select as many perceived barriers as they wished for each technology. “Other” responses were not included.

disregarded subjects who did not respond to at least one question in Q21 and Q22. Logistic regression model results indicated that respondents who worked primarily in Europe were less likely to perceive a high number of barriers to adoption of ATS compared with North Americans [OR=0.56 (95% CI: 0.35, 0.90); $p=0.017$]. Respondents working for all other employer categories were significantly less likely than those working for a national government to indicate barriers (Table 5). Women were significantly less likely than men to indicate barriers, and a 10-y increment in degree year was also negatively associated with barriers. Compared with other sectors, respondents employed by the

pharmaceutical industry also indicated fewer barriers, although the association was not significant.

Drivers of Adoption of Alternative Testing Strategies

We asked the participants to identify which of the several factors, if any, play or will play a significant role in driving the adoption of ATS by them or by their organization. The leading perceived drivers of adoption, for the most part, reflected the normative and methodological factors that have driven the 3Rs concept (replacement, reduction, and refinement) for decades as well as the 2007

Table 5. Logistic regression: Barriers.

Characteristic	OR (95% CI)	Natural log	SE	p-Value
Gender (F) ^a	0.71 (0.52, 0.97)	-0.34	0.16	0.030
Degree year/10 y	0.97 (0.96, 0.99)	-0.03	0.01	0.001
Region (Europe)	0.56 (0.35, 0.90)	-0.58	0.24	0.017
Region (other)	0.89 (0.54, 1.47)	-0.12	0.26	0.635
Employer (small/medium business)	0.36 (0.18, 0.70)	-1.03	0.34	0.003
Employer (large business)	0.60 (0.39, 0.92)	-0.51	0.22	0.021
Employer (state government)	0.28 (0.11, 0.72)	-1.26	0.48	0.008
Employer (academia)	0.31 (0.20, 0.49)	-1.17	0.24	0.001
Employer (other)	0.38 (0.25, 0.59)	-0.96	0.22	0.001
Sector (Pharma)	0.80 (0.55, 1.15)	-0.23	0.19	0.224

Note: Perceived barriers explained by respondents' characteristics compared with a reference group of U.S. males working for a national government organization and not associated with the pharmaceutical industry. CI, confidence interval; OR, odds ratio; SE, standard error.

^aThree subjects classified as "Other" are not included in this analysis.

NAS report (National Research Council 2007), including the need for expedited toxicity information, reducing testing costs, and ethical concerns regarding animal impacts (Figure 2). The third most cited potential driver—demand by regulatory agencies—focuses instead upon the institutional context. The survey data were also noteworthy in what respondents did not commonly identify as significant drivers. For example, the need to compete with others in the industry and demand by the general public, by customers, or by NGOs represented the bottom four answers by overall ranking.

The rank order of drivers changed little when examining the six different types of ATS; it also varied little when compared by age (i.e., degree year), by gender, by European versus North American location, or by employer. Hierarchical clustering generated two groups of respondents based on survey questions about factors potentially playing a significant role in driving the adoption of ATS (Q13), with one group identifying only a limited number of drivers (50% of responders identified between 6 and 18 drivers), and the other group identifying a significantly large number (50% of responders identified between 21 and 36 drivers). As before, in our analysis, we disregarded subjects who did

Table 6. Logistic regression: Drivers.

Characteristic	OR (95% CI)	Natural log	SE	p-Value
Gender (F) ^a	0.75 (0.54, 1.04)	-0.29	0.17	0.085
Degree year/10 y	0.99 (0.97, 1.00)	-0.01	0.01	0.084
Region (Europe)	0.64 (0.40, 1.01)	-0.45	0.23	0.055
Region (other)	0.45 (0.26, 0.79)	-0.79	0.28	0.005
Employer (small/medium business)	0.42 (0.20, 0.87)	-0.87	0.37	0.019
Employer (large business)	0.97 (0.60, 1.58)	-0.03	0.25	0.909
Employer (state government)	0.48 (0.19, 1.21)	-0.74	0.48	0.120
Employer (academia)	0.71 (0.43, 1.18)	-0.34	0.26	0.187
Employer (other)	0.82 (0.51, 1.32)	-0.20	0.25	0.406
Sector (Pharma)	1.19 (0.80, 1.77)	0.18	0.20	0.385

Note: Perceived drivers explained by respondents' characteristics compared with a reference group of U.S. males working for a national government organization and not associated with the pharmaceutical industry. CI, confidence interval; OR, odds ratio; SE, standard error.

^aThree subjects classified as "Other" are not included in this analysis.

not respond to at least one question in Q13. A logistic regression analysis confirmed that no significant differences were observed between groups aside from non-North American/European respondents and from small/medium business employees, who overall perceived fewer drivers to the field when compared with the reference group (Table 6).

Discussion

In this study, we aimed to assess the dynamics and tensions that have animated the evolution of the field of toxicology over the last 10 y. Findings from our survey of 1,381 toxicologists and people working in related areas suggest that current use and perceptions regarding the current viability of novel testing methods (which we refer to, in combination, as "acceptance" of ATS) varied depending on the type of testing approach and depending on the particular application. Our study also revealed that the perceived barriers to and drivers of the adoption of alternative methods differed by geographical location and by type of employer. By contrast, the top drivers, that is to say, expedited information, reduced costs, and regulatory demand, were more commonly shared among respondents. Taken together, these results shed

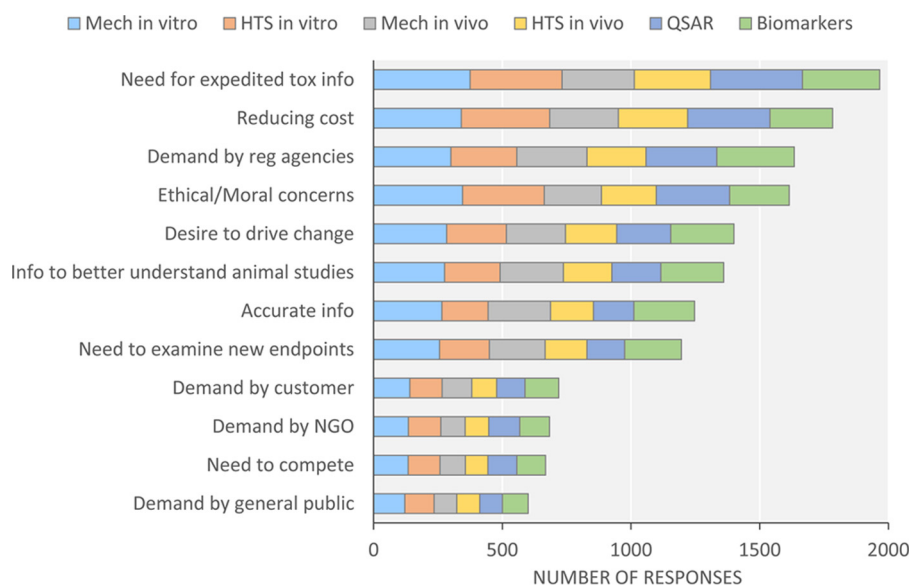


Figure 2. Drivers of adoption of alternative technologies. Respondents were asked to identify which factors, if any, they think play a significant role in driving the adoption of the listed alternative approaches by them or by others in their organization (Question 23). The list of drivers was further divided according to each category of alternative technology. The y-axis represents the total number of responses collected. Respondents were able to select as many perceived drivers as they wished for each technology. "Other" responses were not included.

light on differences among those working in the field regarding the application of ATS and provide some guidance on what is—and what is not—needed to drive integration.

Our survey and analysis had several limitations that should be considered when interpreting our findings. First, although the participants appeared to be representative of SoT members as a whole, our study population was a sample of people who voluntarily responded to survey invitations sent primarily to members of professional societies; therefore, they may not be representative of all scientists and nonscientists who are trained in, working in, or closely linked to the field of toxicology. In addition, the majority of the participants were from the United States, and nearly half were working in the pharmaceutical sector, whereas the numbers of participants from other countries and from some industry sectors and employer types were limited. Numbers of responses varied among the different questions, and it is not possible to determine the underlying causes of missing data in all cases.

The findings with regard to the wide acceptance of ATS in the survey for screening and prioritization is consistent with those of the NAS report and with much of the academic literature (Hartung 2009; National Research Council 2007; Thomas et al. 2013). The NAS report envisioned full integration of ATS into toxicology as a stepwise, decades-long process, with early applications of ATS for screening and prioritization being followed by use in risk assessment (Firestone et al. 2010; National Research Council 2007). Much of the relevant literature also identifies screening and prioritization as appropriate current applications for ATS (Cote et al. 2016; Judson et al. 2010). Views on a more limited role for most ATS methods in QRA in the near term have been expressed (Cote et al. 2016; Crump et al. 2010; Wignall et al. 2014), although some reports have also been more optimistic (Adeleye et al. 2015; Judson et al. 2011). The survey results suggest a divergence from this vision, particularly with respect to the use of ATS in QRA. Depending upon the ATS approach, 25.5% to 53.5% of respondents already view ATS as currently viable for identifying NOAELs and similar levels in QRA, and 5.7–21.2% of respondents already use ATS for this purpose (Table 2). In contrast, and with one exception, a number of respondents (14.7–35.5%) indicated that ATS has no such role to play in the foreseeable future (Table 3). Although further inquiry will be necessary to dissect the root of the differences in perception, the survey data identify potential factors. As discussed above, one factor is familiarity: the more familiar a respondent was with a given technology, the more likely the respondent was to perceive it as currently viable for at least one use (see Tables S4, S5). This finding suggests that an emphasis on education and training may be helpful in enhancing the acceptance and implementation of ATS.

The majority of respondents also indicated that they currently used or that they viewed ATS methods as currently viable for comparative assessments of alternative chemicals, products, or processes (53–69%, depending on the ATS). Alternatives analysis or alternatives assessment is an emerging method used in regulatory and private contexts to identify and select safer alternatives for chemicals of concern in consumer products and industrial processes (National Research Council 2014). Regulatory programs in California and in the European Union mandate that manufacturers or other responsible parties perform alternatives analysis for specified uses of chemicals of particular concern (DTSC 2013; European Parliament and Council 2006). Both programs explicitly allow for the use of ATS in comparing the risks of the incumbent chemical with potential alternatives (ECHA 2011; OEHHA 2012). The perceived viability of ATS for these uses may at least partly reflect the comparative orientation of alternatives analysis; finely grained, quantitative results specifying a single acceptable

exposure level are less critical in this context than in quantitative risk assessment (Hjorth et al 2017; Niska et al. 2008).

Regarding barriers and drivers of adoption, one striking paradox emerges. Concern about regulatory acceptance was the most commonly selected factor that “plays a significant role in inhibiting the adoption of ATS” by respondents or by others in their organizations, but demand by regulatory agencies was also frequently cited as a factor that “plays (or will play) a significant role in driving the adoption” of ATS by respondents or by other people in their organizations. Cross-tabulations of the questions about barriers and drivers indicated that that more than half of the respondents who cited regulatory acceptance as a current barrier also cited regulatory demand as a current or future driver of ATS adoption (data not shown). This paradox may reflect the nature of the regulatory enterprise itself, in which regulatory agencies must balance the need for cost-effective, timely toxicity information against the understandable importance placed by them upon legitimacy, consistency in regulatory action, and cautious appraisal and considered adoption of new technologies in order to avoid the perception of rash agency action. We believe that the use of ATS data for regulatory purposes will become increasingly common over time.

General resistance to change was the sixth most commonly cited factor (out of 13) that “plays a significant role in inhibiting the adoption” of ATS. Once established, technical capacities and routine practices of the sort seen in conventional regulatory toxicology are difficult to change because of the resources and incentive structures that develop over time to support them (Malloy 2011). Of course, business firms and regulatory agencies are not monolithic; in some cases, specific individuals and groups within the organization may be more focused upon generating and deploying new methods and practices. For example, innovation is central to the mission of the Office of Research and Development within the U.S. EPA; accordingly, this office has been spearheading advances in ATS at the U.S. EPA (Judson et al. 2010). However, real change within regulatory programs may necessitate mandating the use of ATS approaches where scientifically appropriate. Similarly, changes to conventional validation processes will likely be necessary to accelerate the availability and credibility of emerging ATS methods.

Finally, it should be noted that both barriers and drivers are dynamic and interactive. For example, in line with the perceived importance of demands by regulatory agencies as a driver to the adoption of ATS, recent changes in U.S. chemicals law could have a significant impact. In June 2016, the Frank R. Lautenberg Chemical Safety for the 21st Century Act (the “Act”) was enacted, substantially reforming the federal Toxic Substances Control Act (TSCA) (U.S. Congress 2016). Among other things, the Act amends the TSCA to include three provisions focused upon the reduction of testing on vertebrates. The first provision creates a “soft” mandate for the use of ATS in screening. The provision requires that before testing on vertebrate animals, the U.S. EPA must consider reasonably available information such as computational toxicology and bioinformatics, HTS methods, and the prediction models of those methods to the extent that they are practicable and scientifically justified. The second provision directs the U.S. EPA to encourage and facilitate the use of scientifically valid alternative test methods and chemical grouping approaches to reduce or replace testing on vertebrate animals. The third provision requires the U.S. EPA to develop a strategic plan by June 2018 to promote development and implementation of alternative test strategies, including computational toxicology and bioinformatics, HTS screening, testing of categories of chemicals, *in vitro* studies, and systems biology. On their face, the three provisions strongly encourage the development and use of ATS within the TSCA program, yet they do little to address the fifth- and sixth-ranked social/legal/institutional

barriers, general resistance to change and complexity/slowness of the Interagency Coordinating Committee on the Validation of Alternative Methods/European Centre for the Validation of Alternative Methods (ICCVAM/ECVAM) validation processes, respectively. Regarding resistance to change, which may be present within regulated entities and within the U.S. EPA as well, the provisions permit and encourage the use of ATS where scientifically appropriate, but, with very limited exceptions, they do not mandate ATS use in such circumstances. Additionally, the statute requires no efforts by relevant agencies to streamline the ICCVAM validation process.

Other provisions of the TSCA reform legislation may create regulatory demand sufficient to accelerate the adoption of ATS approaches in regulatory decision making. Under the statute, the U.S. EPA faces challenging deadlines for screening and prioritization of existing chemicals in commerce and for risk evaluation of those chemicals ultimately prioritized for review and action. As a matter of necessity, regulated entities and the agency itself may come to increasingly rely upon ATS approaches to meet their new obligations. Even without explicit mandates to adopt appropriate ATS methods, this regulatory demand may indirectly overcome inertia and other institutional constraints on the adoption of ATS.

Conclusion

The present survey was conducted to determine the extent of current ATS use and to gather information about barriers and drivers that might help ATS proponents increase its acceptance. Overall, it appears that toxicology as a field has followed the path described by the NAS report in 2007, with initial use of ATS for screening and prioritization, eventually followed by use for risk assessment (National Research Council 2007). However, the respondents differed with regard to whether ATS was currently viable, versus unlikely to be viable for the foreseeable future, to assess end points in their primary area of interest or concern, particularly with respect to QRA. The perceived obstacles to further development and adoption were largely common to all forms of ATS, as were the factors that participants cited as current or future drivers. In addition to technological and validation issues, institutional factors relating to regulatory acceptance, resistance to change, and constraints associated with conventional validation processes are viewed as significant barriers. Respondents indicated that along with demand by regulatory agencies, the need for expedited toxicity information, the need for reduced toxicity testing costs, and ethical and moral concerns regarding the use of animals in testing were the four most commonly cited factors that played or will play a significant role in their own or in their organization's adoption of ATS, whereas the need to compete with others in their industry and demand by customers, NGOs, and the general public were the four factors least likely to be cited as current or future drivers of ATS adoption. In conclusion, taking into account the aforementioned survey limitations, the survey findings provide important information that may be used to inform efforts to promote the use and acceptance of ATS. Based on our interpretation of the findings, we recommend that in addition to continued scientific research and development, a two-pronged intervention be considered to move the field of toxicology further into the 21st century; this intervention would consist of bottom-up coordinated efforts from stakeholders to encourage adoption of ATS approaches coupled with top-down legal and institutional changes focused on expediting regulatory acceptance of a diverse range of technologies. For example, increased efforts by interested nongovernmental organizations to encourage the development and adoption of ATS by regulators and businesses and to educate consumers and the general public about the potential of ATS could enhance adoption. Similarly,

sustained efforts by change agents within federal institutions such as the Interagency Center for the Evaluation of Alternative Toxicological Methods could reduce barriers to validation, particularly in light of recent changes to the TSCA (Casey 2016).

Acknowledgments

The authors would like to thank S. Shafer for his assistance with statistical analysis and B. Nyblade for his comments on the manuscript. We are also especially grateful to the Society of Toxicology leadership and B. Eidemiller for their help with disseminating the survey. This work was supported by a Transdisciplinary Seed Grant to T.M. and P.A. from the Office of the Vice Chancellor for Research at the University of California, Los Angeles, and by an Environmental Solutions Venture Grant from the Institute of the Environment and Sustainability at the University of California, Los Angeles.

References

- Adeleye Y, Andersen M, Clewell R, Davies M, Dent M, Edwards S, et al. 2015. Implementing toxicity testing in the 21st century (TT21C): making safety decisions using toxicity pathways, and progress in a prototype risk assessment. *Toxicology* 332:102–111, PMID: 24582757, <https://doi.org/10.1016/j.tox.2014.02.007>.
- Andersen ME, Krewski D. 2010. The vision of toxicity testing in the 21st century: Moving from discussion to action. *Toxicol Sci* 117(1):17–24, PMID: 20573784, <https://doi.org/10.1093/toxsci/kfq188>.
- Browne P, Judson RS, Casey WM, Kleinstreuer NC, Thomas RS. 2015. Screening chemicals for estrogen receptor bioactivity using a computational model. *Environ Sci Technol* 49(14):8804–8814, PMID: 26066997, <https://doi.org/10.1021/acs.est.5b02641>.
- Casey WM. 2016. Advances in the development and validation of test methods in the united states. *Toxicol Res* 32(1):9–14, PMID: 26977254, <https://doi.org/10.5487/TR.2016.32.1.009>.
- Collins FS, Gray GM, Bucher JR. 2008. Toxicology. Transforming environmental health protection. *Science* 319(5865):906–907, PMID: 18276874, <https://doi.org/10.1126/science.1154619>.
- Cote I, Andersen ME, Ankley GT, Barone S, Birnbaum LS, Boekelheide K, et al. 2016. The next generation of risk assessment multi-year study-highlights of findings, applications to risk assessment, and future directions. *Environ Health Perspect* 124(11):1671–1682, PMID: 27091369, <https://doi.org/10.1289/EHP233>.
- Crump KS, Chen C, Louis TA. 2010. The future use of *in vitro* data in risk assessment to set human exposure standards: Challenging problems and familiar solutions. *Environ Health Perspect* 118(10):1350–1354, PMID: 20562051, <https://doi.org/10.1289/ehp.1001931>.
- DTSC (California Department of Toxic Substances Control). 2013. Safer consumer product regulations. <http://www.dtsc.ca.gov/SCPRRegulations.cfm> [accessed 17 July 2017].
- ECHA (European Chemicals Agency). 2011. *Guidance on the Preparation of an Application for Authorisation*. Helsinki, Finland:ECHA.
- European Parliament and Council. 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing 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, 93/67/EEC, 93/105/EC and 2000/21/EC. *OJ L* 49(396):1–849.
- Firestone M, Kavlock R, Zenick H, Kramer M, U.S. Environmental Protection Agency Working Group on the Future of Toxicity Testing. 2010. The U.S. Environmental Protection Agency strategic plan for evaluating the toxicity of chemicals. *J Toxicol Environ Health B Crit Rev* 13(2–4):139–162, PMID: 20574895, <https://doi.org/10.1080/10937404.2010.483178>.
- Gomez A. (U.S. Government Accountability Office) 2013. Chemical regulation: Observations on the Toxic Substances Control Act and EPA implementation, statement of Alfredo Gomez. <https://www.gao.gov/assets/660/655202.pdf> [accessed 17 July 2017].
- Hartung T. 2009. A toxicology for the 21st century—mapping the road ahead. *Toxicol Sci* 109(1):18–23, PMID: 19357069, <https://doi.org/10.1093/toxsci/kfp059>.
- Hjorth R, Hansen SF, Jacobs M, Tickner J, Ellenbecker M, Baun A. 2017. The applicability of chemical alternatives assessment for engineered nanomaterials. *Integr Environ Assess Manag* 13(1):177–187, PMID: 26887668, <https://doi.org/10.1002/ieam.1762>.

- Judson R, Richard A, Dix DJ, Houck K, Martin M, Kavlock R, et al. 2009. The toxicity data landscape for environmental chemicals. *Environ Health Perspect* 117(5):685–695, PMID: 19479008, <https://doi.org/10.1289/ehp.0800168>.
- Judson RS, Houck KA, Kavlock RJ, Knudsen TB, Martin MT, Mortensen HM, et al. 2010. *In vitro* screening of environmental chemicals for targeted testing prioritization: The ToxCast project. *Environ Health Perspect* 118(4):485–492, PMID: 20368123, <https://doi.org/10.1289/ehp.0901392>.
- Judson RS, Kavlock RJ, Setzer RW, Hubal EA, Martin MT, Knudsen TB, et al. 2011. Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chem Res Toxicol* 24(4):451–462, PMID: 21384849, <https://doi.org/10.1021/tx100428e>.
- Leist M, Hartung T. 2013. Inflammatory findings on species extrapolations: Humans are definitely no 70-kg mice. *ALTEX* 30(2):227–230, PMID: 23665808.
- Leung MC, Williams PL, Benedetto A, Au C, Helmcke KJ, Aschner M, et al. 2008. *Caenorhabditis elegans*: An emerging model in biomedical and environmental toxicology. *Toxicol Sci* 106(1):5–28, PMID: 18566021, <https://doi.org/10.1093/toxsci/kfn121>.
- Liaw A, Wiener M. 2002. Classification and regression by random forest. *R News* 2:18–22.
- Malloy TF. 2011. Nanotechnology regulation: A study in claims making. *ACS Nano* 5(1):5–12, PMID: 21261305, <https://doi.org/10.1021/nn103480e>.
- National Research Council. 2007. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington, DC: The National Academies Press.
- National Research Council. 2014. *A Framework to Guide the Selection of Chemical Alternatives*. Washington, DC: The National Academies Press.
- Neltner TG, Alger HM, Leonard JE, Maffini MV. 2013. Data gaps in toxicity testing of chemicals allowed in food in the United States. *Reprod Toxicol* 42:85–94, PMID: 23954440, <https://doi.org/10.1016/j.reprotox.2013.07.023>.
- Niska H, Tuppurainen K, Skon JP, Mallett AK, Kolehmainen M. 2008. Characterisation of the chemical and biological properties of molecules with QSAR/QSPR and chemical grouping, and its application to a group of alkyl ethers. *SAR QSAR Environ Res* 19(3–4):263–284, PMID: 18484498, <https://doi.org/10.1080/10629360802085033>.
- OECD (Organisation for Economic Co-operation and Development). 2005. Development, validation and regulatory acceptance of new and updated internationally acceptable test methods in hazard assessment. [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=en/jm/mono\(2005\)14](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?doclanguage=en&cote=en/jm/mono(2005)14) [accessed 17 July 2017].
- OEHHA (Office of Environmental Health Hazard Assessment). 2012. Green chemistry hazard traits, toxicological and environmental endpoints and other relevant data. [https://govt.westlaw.com/calregs/Browse/Home/California/CaliforniaCodeofRegulations?guid=I6E0E45C032A411E186A4EF11E7983D17&originationContext=documenttoc&transitionType=Default&contextData=\(sc.Default\)](https://govt.westlaw.com/calregs/Browse/Home/California/CaliforniaCodeofRegulations?guid=I6E0E45C032A411E186A4EF11E7983D17&originationContext=documenttoc&transitionType=Default&contextData=(sc.Default)) [accessed 17 July 2017].
- S. Congress. 2016. Frank R. Lautenberg Chemical Safety for the 21st Century Act. Pub L 114–182 (114th Congress, 22 June 2016). <https://www.congress.gov/114/plaws/publ182/PLAW-114publ182.pdf> [accessed 17 July 2017].
- Stephenson J. (U. S. Government Accountability Office) 2009. Chemical regulation: Observations on improving the toxic substances control act, statement of john stephenson. <http://www.gao.gov/assets/130/123792.pdf> [accessed 17 July 2017].
- Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, et al. 2013. Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. *Toxicol Sci* 136(1):4–18, PMID: 23958734, <https://doi.org/10.1093/toxsci/kft178>.
- Tice RR, Austin CP, Kavlock RJ, Bucher JR. 2013. Improving the human hazard characterization of chemicals: A Tox21 update. *Environ Health Perspect* 121(7):756–765, PMID: 23603828, <https://doi.org/10.1289/ehp.1205784>.
- Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, et al. 2014. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. *Environ Health Perspect* 122(5):499–505, PMID: 24569956, <https://doi.org/10.1289/ehp.1307539>.
- Yang L, Ho NY, Alshut R, Legradi J, Weiss C, Reischl M, et al. 2009. Zebrafish embryos as models for embryotoxic and teratological effects of chemicals. *Reprod Toxicol* 28(2):245–253, PMID: 19406227, <https://doi.org/10.1016/j.reprotox.2009.04.013>.