Message

From: Faeth, Lisa [Faeth.Lisa@epa.gov]
Sent: 6/7/2018 3:17:28 PM
To: Askinazi, Valerie [Askinazi.Valerie@epa.gov]; Barkas, Jessica [barkas.jessica@epa.gov]; Beck, Nancy [Beck.Nancy@epa.gov]; Bertrand, Charlotte [Bertrand.Charlotte@epa.gov]; Blair, Susanna [Blair.Susanna@epa.gov]; Blunck, Christopher [Blunck.Chris@epa.gov]; Brown, Sam [Brown.Sam@epa.gov]; Buster, Pamela [Buster.Pamela@epa.gov]; Canavan, Sheila [Canavan.Sheila@epa.gov]; Caraballo, Mario [Caraballo.Mario@epa.gov]; Carroll, Megan [Carroll.Megan@epa.gov]; Cherepy, Andrea [Cherepy.Andrea@epa.gov]; Christian, Myrta [Christian.Myrta@epa.gov]; Corado, Ana [Corado.Ana@epa.gov]; Davies, Clive [Davies.Clive@epa.gov]; DeDora, Caroline [DeDora.Caroline@epa.gov]; Devito, Steve [Devito.Steve@epa.gov]; Doa, Maria [Doa.Maria@epa.gov]; Drewes, Scott [Drewes.Scott@epa.gov]; Dunton, Cheryl [Dunton.Cheryl@epa.gov]; Ebzery, Joan [Ebzery.Joan@epa.gov]; Edelstein, Rebecca [Edelstein.Rebecca@epa.gov]; Edmonds, Marc [Edmonds.Marc@epa.gov]; Eglsaer, Kristie [Eglsaer.Kristie@epa.gov]; Elwood, Holly [Elwood.Holly@epa.gov]; Faeth, Lisa [Faeth.Lisa@epa.gov]; Farquharson, Chenise [Farquharson.Chenise@epa.gov]; Fehrenbacher, Cathy [Fehrenbacher.Cathy@epa.gov]; Feustel, Ingrid [feustel.ingrid@epa.gov]; Frank, Donald [Frank.Donald@epa.gov]; Gibson, Hugh [Gibson.Hugh@epa.gov]; Gimlin, Peter [Gimlin.Peter@epa.gov]; Gorder, Chris [Gorder.Chris@epa.gov]; Gordon, Brittney [Gordon.Brittney@epa.gov]; Grant, Brian [Grant.Brian@epa.gov]; Gray, Shawna [Gray.Shawna@epa.gov]; Groeneveld, Thomas [Groeneveld.Thomas@epa.gov]; Guthrie, Christina [Guthrie.Christina@epa.gov]; Helfgott, Daniel [Helfgott.Daniel@epa.gov]; Henry, Tala [Henry.Tala@epa.gov]; Kapust, Edna [Kapust.Edna@epa.gov]; Kemme, Sara [kemme.sara@epa.gov]; Koch, Erin [Koch.Erin@epa.gov]; Krasnic, Toni [krasnic.toni@epa.gov]; Lavoie, Emma [Lavoie.Emma@epa.gov]; Leczynski, Barbara [leczynski.barbara@epa.gov]; Lee, Mari [Lee.Mari@epa.gov]; Lee, Virginia [Lee.Virginia@epa.gov]; Leopard, Matthew (OEI) [Leopard.Matthew@epa.gov]; Liva, Aakruti [Liva.Aakruti@epa.gov]; Lobar, Bryan [Lobar.Bryan@epa.gov]; Mclean, Kevin [Mclean.Kevin@epa.gov]; Menasche, Claudia [Menasche.Claudia@epa.gov]; Morris, Jeff [Morris.Jeff@epa.gov]; Moss, Kenneth [Moss.Kenneth@epa.gov]; Mottley, Tanya [Mottley.Tanya@epa.gov]; Moyer, Adam [moyer.adam@epa.gov]; Myers, Irina [Myers.Irina@epa.gov]; Myrick, Pamela [Myrick.Pamela@epa.gov]; Nazef, Laura [Nazef.Laura@epa.gov]; Ortiz, Julia [Ortiz.Julia@epa.gov]; Owen, Elise [Owen.Elise@epa.gov]; Parsons, Doug [Parsons.Douglas@epa.gov]; Passe, Loraine [Passe.Loraine@epa.gov]; Pierce, Alison [Pierce.Alison@epa.gov]; Pratt, Johnk [Pratt.Johnk@epa.gov]; Price, Michelle [Price.Michelle@epa.gov]; Reese, Recie [Reese.Recie@epa.gov]; Reisman, Larry [Reisman.Larry@epa.gov]; Rice, Cody [Rice.Cody@epa.gov]; Richardson, Vickie [Richardson.Vickie@epa.gov]; Ross, Philip [Ross.Philip@epa.gov]; Sadowsky, Don [Sadowsky.Don@epa.gov]; Santacroce, Jeffrey [Santacroce.Jeffrey@epa.gov]; Saxton, Dion [Saxton.Dion@epa.gov]; Scarano, Louis [Scarano.Louis@epa.gov]; Scheifele, Hans [Scheifele.Hans@epa.gov]; Schmit, Ryan [schmit.ryan@epa.gov]; Schweer, Greg [Schweer.Greg@epa.gov]; Selby-Mohamadu, Yvette [Selby-Mohamadu.Yvette@epa.gov]; Seltzer, Mark [Seltzer.Mark@epa.gov]; Sheehan, Eileen [Sheehan.Eileen@epa.gov]; Sherlock, Scott [Sherlock.Scott@epa.gov]; Simons, Andrew [Simons.Andrew@epa.gov]; Sirmons, Chandler [Sirmons.Chandler@epa.gov]; Slotnick, Sue [Slotnick.Sue@epa.gov]; Smith, David G. [Smith.DavidG@epa.gov]; Stedeford, Todd [Stedeford.Todd@epa.gov]; Strauss, Linda [Strauss.Linda@epa.gov]; Symmes, Brian [Symmes.Brian@epa.gov]; Tanner, Barbara [Tanner.Barbara@epa.gov]; Thompson, Tony [Thompson.Tony@epa.gov]; Tierney, Meghan [Tierney.Meghan@epa.gov]; Tillman, Thomas [Tillman.Thomas@epa.gov]; Tomassoni, Guy [Tomassoni.Guy@epa.gov]; Tran, Chi [Tran.Chi@epa.gov]; Vendinello, Lynn [Vendinello.Lynn@epa.gov]; Wallace, Ryan [Wallace.Ryan@epa.gov]; Wheeler, Cindy [Wheeler.Cindy@epa.gov]; Widawsky, David [Widawsky.David@epa.gov]; Williams, Aresia [Williams.Aresia@epa.gov]; Williams, Bridget [Williams.Bridget@epa.gov]; Williamson, Tracy [Williamson.Tracy@epa.gov]; Wills, Jennifer [Wills.Jennifer@epa.gov]; Wise, Louise [Wise.Louise@epa.gov]; Wolf, Joel [Wolf.Joel@epa.gov]; Wright, Tracy [Wright.Tracy@epa.gov]; Yowell, John [yowell.john@epa.gov]
Subject: News Articles (For EPA Distribution Only)

BNA DAILY ENVIRONMENT REPORT ARTICLES

Two Pruitt Aides Who Got Hefty Raises Resign

By David Schultz (Bloomberg Environment), Jennifer A. Dlouhy (Bloomberg News) and Jennifer Jacobs (Bloomberg News)

Posted June 6, 2018, 6:35 PM

Two senior aides to EPA Administrator Scott Pruitt are resigning, with one of them stepping down after the revelation that she performed personal tasks for Pruitt, including exploring buying a used mattress from one of President Donald Trump's hotels.

Bill With Modest Cut to EPA Funding Approved by House Panel

By Dean Scott

Posted June 6, 2018, 5:03 PM

The House Appropriations Committee rejected the White House's proposed 25 percent slash of the Environmental Protection Agency's budget, but did vote to cut the agency's spending by about $100 million under a bill approved June 6.

What's in Your Product? N.Y. to Require Chemical Disclosure

By Gerald B. Silverman

Posted June 6, 2018, 3:19 PM

S.C. Johnson & Son Inc., Unilever, Procter & Gamble Co., and other manufacturers of household cleaners will have just over a year before they must disclose the chemical ingredients and other information about their products in New York state.

INSIDEEPA.COM ARTICLES

GOP Criticisms Of Pruitt's Scandals Mount, Bolstering Democratic Attacks

Some key Republicans are intensifying their criticisms of EPA Administrator Scott Pruitt's ethics and spending in the wake of new reports showing he directed subordinates to perform allegedly unlawful personal tasks for him, bolstering Democrats' efforts to use Pruitt as a prime example in their political campaign against the Trump administration's "corruption."

EPA Faces Calls To Overhaul TSCA Test Method Plan On Eve Of Release

As EPA prepares to unveil a congressionally mandated plan for how to reduce animal testing of chemicals, the agency is facing competing calls from environmentalists and animal rights groups who are still at odds over how quickly officials can end or limit such testing.

GREENWIRE ARTICLES

Emails show stressed staffers ghostwriting op-ed for Pruitt

Maxine Joselow, E&E News reporter

Published: Wednesday, June 6, 2018

Samantha Dravis arriving at Trump Tower with Scott Pruitt in December 2016. Andrew Harnik/Associated Press

The pressure was on last spring as EPA political appointees assembled an op-ed for Administrator Scott Pruitt on revising clean car rules.

"Please kill me," Samantha Dravis, who was then Pruitt's policy chief, wrote in an email to her colleagues. "I can't get used to this Dell laptop. My track changes are not tracking ... did you guys lock editing on this or something??"

In the works on that day in March 2017 was a *USA Today* op-ed. The rush was on, spurred by an email from EPA Chief of Staff Ryan Jackson.

"Guys please hurry and get this get this [sic] placed," he wrote. "I was assuming that was done well before now."

https://www.eenews.net/greenwire/2018/06/06/stories/1060083669

Trump has rolled back, delayed 1,500-plus rules — watchdog

Maxine Joselow, E&E News reporter

Published: Wednesday, June 6, 2018

The Trump administration has withdrawn or delayed more than 1,500 rules — much more than the number under President Obama but roughly parallel with the number under President George W. Bush, according to a report from a left-leaning watchdog group released today.

The report from Public Citizen looked at 20 years of rulemakings reported by the White House Office of Management and Budget in the so-called Unified Agenda, an administration's sweeping regulatory plan.

It found that federal agencies under Trump have withdrawn 780 rulemakings — an 85 percent increase from under Obama and a 3 percent decrease from under Bush.

Of those withdrawn rulemakings, 279 were deemed "significant," meaning they had an annual effect on the economy of more than $100 million.

https://www.eenews.net/greenwire/2018/06/06/stories/1060083681

Country to ban single-use plastic by 2022

Published: Wednesday, June 6, 2018

India will do away with all single-use plastic by 2022, Prime Minister Narendra Modi said yesterday.

The move would be drastic in one of the largest countries — and fastest-growing economies — in the world. It would be the most sweeping ban of the 60 nations worldwide that have moved to reduce plastic pollution.

"The choices that we make today will define our collective future," Modi said. "The choices may not be easy. But through awareness, technology, and a genuine global partnership, I am sure we can make the right choices."

https://www.eenews.net/greenwire/2018/06/06/stories/1060083637

CHEMICAL WATCH ARTICLES

Colombia publishes draft decree on industrial chemical management

Ministries given a year to set up registry

6 June 2018 / Colombia, GHS

Colombia has published a draft decree on regulating the management of industrial use chemicals.

The draft is a revised version of one released in February last year for public consultation. It comes out of the country's national policy for risk management related to the use of chemical substances. This was established to help Colombia comply with the requirements of OECD membership, which it was formally invited to join on 25 May.

The decree says its main purpose is the protection of health and the environment. Among its 28 articles, the draft sets out three principal ways of achieving this. They are:

- registration of chemicals;
- risk assessment; and
- risk management and reduction.

Three ministries – commerce, industry and tourism (MinCIT), health and social protection (MinSalud) and environment and sustainable development (Minambiente) – have been given a year to establish a chemicals registry regulation. This dates from the decree's publication in the country's *Official Gazette*.

Risk assessment and risk management processes must be in place two years after the registry is up and running. The proposal says manufacturers or importers of substances identified as carcinogenic, mutagenic or reprotoxic (CMR), or with properties of persistence or bioaccumulation, or which present a short-term hazard to the environment, will have to prepare such reduction or management programmes.

The manufacturer or importer must, the draft says, "make an annual report on the execution of the actions proposed in the risk reduction and management programmes for the environment or for health".

Among its stipulations, the draft lays down the obligations of:

- importers and manufacturers;
- marketers and distributors;
- transporters; and
- users.

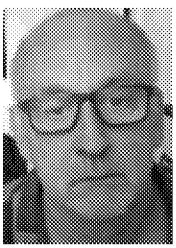
A committee is being created to oversee the implementation of the regulation. The Intersectoral Committee for the Follow-up of the Management of Chemical Substances for Industrial Use (Cisquin) will have representatives from each of the key ministries mentioned, plus the labour ministry (MinTrabajo).

The decree refers to the sixth edition of the Globally Harmonized System (GHS) of classification and labelling of chemicals to identify hazardous substances. This was also used as a reference for the GHS draft Decree – notified to the WTO last November – which has not yet been published in the *Official Gazette*.

An annex to the decree details hazard categories that would be required in a registry of substances.

The decree gives no date of its entry into force but will follow publication in the *Official Gazette*. However, the WTO has been notified of the proposal and is carrying out a consultation until 22 July.

Note: Your access to this subscriber-only article is through a corporate subscription



Nick Hazlewood

News editor

Related Articles

- Colombia announces intention to adopt GHS

Further Information:

- Draft decree

Turkish ministry approves seven KKDIK expert training institutes

7 June 2018 / Turkey

The number of Turkish institutions authorised to train chemical assessment experts under the country's KKDIK law has risen to seven, according to the Ministry of Environment and Urbanisation (MoEU).

As part of the REACH-like law, Turkey requires such professionals – KDUs in Turkish – to be trained and accredited by government-approved agencies before they can sign chemical safety reports (CSRs).

KKDIK has an extra annex – Annex 18 – specifying the qualifications necessary to become a KDU. They must undergo at least 64 hours of face-to-face training and pass an exam before they can be issued with a certificate from an accredited organisation.

All companies doing business in the chemicals industry in Turkey, including foreign multinationals, must employ trained KDUs.

According to information on the ministry's website the organisations providing KDU training are:

- Cetin Akademi consultancy – Tekirdag;
- Crad consultancy – Istanbul;
- DATA consultancy – Konya;
- DORUK consultancy – Gebze
- TMMOB - Chamber of Chemical Engineers – Ankara;
- Turkish Chemical Manufacturers Association (TKSD) – Istanbul;
- TUV Austria Certification and Training – Ankara.

TUV Austria was previously authorised under the name NBC certification and training, but the company has changed its name following a partnership with TUV Austria Group.

Related Articles

- Turkey publishes law modelled on REACH
- Race begins in Turkey for KKDIK expert training

- Turkey KKDIK expert training initiatives gain pace

Further Information:

- List of consultancies
- KKDIK annex 18
- KKDIK Regulation

Snur would ban new use of asbestos in the US

EPA affirms plan to ignore 'legacy' asbestos in risk evaluation

7 June 2018 / Built environment, Chemical manufacturing, TSCA, United States

The US EPA has announced it will propose a significant new use rule (Snur) to prevent industry from freely initiating new uses of asbestos, or resuming applications abandoned after the substance's connection to mesothelioma and lung cancer was established.

The agency's move comes even as it holds on to its plan to exclude 'legacy' materials from the risk evaluation of asbestos.

The Snur, released 2 June in pre-publication form, would apply to building materials other than cement. It also covers sealants, adhesives, electrical paper, reinforced plastics and separators in fuel cells, as well as novel uses.

If adopted, it would require industry to notify the EPA of new uses, which it could then regulate or deny.

The proposal is apparently the initiative that EPA Administrator Scott Pruitt referred to in recent appearances on Capitol Hill, where he said legacy asbestos is a "meaningful concern" that was under "active consideration".

Meanwhile, the EPA's 'problem formulation' documents confirm it will consider only ongoing uses in risk evaluation. That plan was set out in June last year in a framework rule and in preliminary scoping documents for each of the first ten chemicals to be evaluated.

The main ongoing use of asbestos is in the chlor-alkali industry. The EPA will also look at:

- sheet gaskets for titanium dioxide production;
- brake blocks used in oil drilling;
- aftermarket automotive products;

- cements; and
- imported textiles.

NGOs unimpressed

The exclusion of abandoned uses means the evaluation will not consider risks posed by old asbestos building materials or their disposal. While not new, this exclusion of 'legacy' asbestos has drawn much of the fire aimed by NGOs at the problem formulations. And they were not impressed by the Snur.

"With these exclusions, it is clear that the EPA will fail to evaluate the full risk of asbestos," an outcome that "represents an irresponsible and unjust rollback" of TSCA reform, said Linda Reinstein, President of the Asbestos Disease Awareness Organization.

"Scott Pruitt's blatant attempt to bamboozle Americas by stating he's taking 'important, unprecedented action on asbestos', while avoiding a ban, is reprehensible," she added.

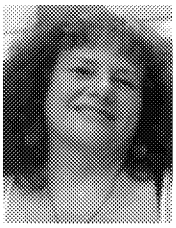
A Snur would prevent companies from deciding to use asbestos once again, which they could now do without notice, Melanie Benesh, legislative attorney at the Environmental Working Group, told Chemical Watch.

"In that sense, it's better than the status quo," she said. "But if these uses have been abandoned and we know asbestos is dangerous and not needed for these applications, why not ban it?"

And Liz Hitchcock, acting director at Safer Chemicals, Healthy Families, told Chemical Watch that a Snur is a "useful stopgap measure", but is not a substitute for regulation.

The direction taken in the problem formulations is troubling, she said. "Asbestos is the poster child [for] a narrowing of scope that will not be protective of public health," she said.

The EPA tried to ban asbestos in 1989, but a federal court threw out the regulation, underlining the weakness of the old TSCA.



Julie Miller

Reporter

Related Articles

- Senate committee cross questions EPA Administrator on chemical policy
- EPA issues TSCA 'problem formulation' documents
- US EPA issues final TSCA framework rules
- US EPA issues final TSCA framework rules

- EPA names first ten chemicals for new TSCA evaluations
- NGOs call for stricter regulations on chlor-alkali industry

Further Information:

- Snur (prepublication)
- Risk assessment documents on asbestos

Industry calls for tiered fees for TSCA risk evaluation

Stakeholders flag up consortia concerns

7 June 2018 / TSCA, United States

Industry groups have called on the US EPA to modify its "relatively simplistic" proposed fees for conducting TSCA risk evaluations. The fees, they say, should more accurately reflect the costs of the assessments.

Issued in February, the proposed fees rule would charge manufacturers $1.35m for EPA-initiated risk evaluations, and either $1.3m or $2.6m for manufacturer-requested risk evaluations, depending on whether the requested substance is a Work Plan chemical or not.

But in response, industry groups say fees should reflect the wide variability of the costs associated with a risk evaluation. These factors include:

- how many uses will be evaluated;
- how long the chemical has been active in commerce;
- the extent to which workers or consumers are exposed; and
- the volume of data that already exists.

The current approach is "unfair to data-rich and uncontroversial chemicals," the National Association of Manufacturers said in its comments on the proposed fees.

And the Ad Hoc Downstream Users Coalition – which comprises paper, motor equipment, toy and tyre manufacturing trade groups – said a 'one-size-fits-all' approach may "lead to unintended consequences like product discontinuation or supply disruption".

The US Chamber of Commerce added that substances produced only by one or two manufacturers could disappear from the marketplace if those producers "opt out of production instead of paying an enormous fee".

In its comments the American Chemistry Council suggested the agency tie its risk evaluation fees to the actual cost it incurs. The ACC proposed setting this price tag at approximately 35% of EPA's actual costs.

Linking the fee to actual costs would "encourage manufacturers to provide as much information as possible to reduce the agency's costs," it added.

Industry groups also roundly objected to the EPA's proposal to require payment of the full risk evaluation fees within 60 days of the agency publishing its final scope of the assessment.

"Requiring upfront payment for a multi-year project is unfair to industry, unnecessary, and contrary to common commercial practices that stagger such payments," wrote B&C Consortia Management (BCCM).

The ACC suggested the agency charge a down payment at the scoping phase, and a final closing payment at the risk evaluation's conclusion.

Consortia challenges

In its proposal, the EPA had said it thought a 60-day timeframe would be "sufficient" for manufacturers to form a consortium and decide on the partial fee payments each member would be responsible for.

But the BCCM said consortia management is "difficult, time consuming, and seldom successful absent extraordinary effort without EPA engagement."

REACH has demonstrated that forming a consortium within a 60-day window is "next to impossible", added the US Chamber.

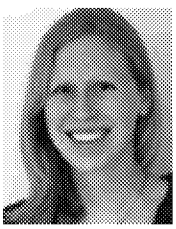
Speciality chemicals group Socma recommended that the EPA help manufacturers in the formation of consortia and "offer logistical support when challenges arise".

And several groups asked that the agency publish a list of companies it knows to be engaged in the manufacture of substances, to assist in identifying consortia participants and ensure that all manufacturers of a substance contribute to its associated fees.

But Socma was among those who warned that relying solely on existing reporting schemes like the chemical data reporting (CDR) rule to identify manufacturers of a substance would lead to an "incomplete list of users".

Instead it proposed a "formal notification system for high priority chemicals [that would] discourage free-rider scenarios and ensure that EPA captures the full universe of manufacturers that have engaged (or will engage) in the manufacture of a risk evaluation chemical."

The ACC also argued that the EPA should consider setting up a mechanism by which late market entrants – or free-riders – could be "held accountable for their fair share of fees, perhaps by establishing reimbursement rules."



Kelly Franklin

North America editor

Related Articles

- US EPA: highest fees for risk evaluation of existing chemicals

Further Information:

- Public docket

US EPA received 17 pre-manufacture notices in February

7 June 2018 / TSCA, United States

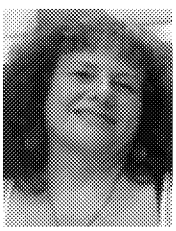
The US EPA received 17 new pre-manufacture notices (PMNs) in February, and recorded ten amendments to existing PMNs.

During January, it also received a new significant new use notice (Snun) and an amendment to an existing Snun.

The agency reported 32 new notices of commencement (NOCs), and received test data on 12 substances.

The EPA reported updates to its February new substance programme in a *Federal Register* notice on 5 June. This is two weeks after separate notices for January, November and December.

Notices for these months mark a new reporting policy. This offers more detail on PMN submissions and testing data received in connection with consent orders.



Julie Miller

Reporter

Related Articles

- US EPA received 20 new PMNs in January
- US EPA received 38 new PMNs in November and December

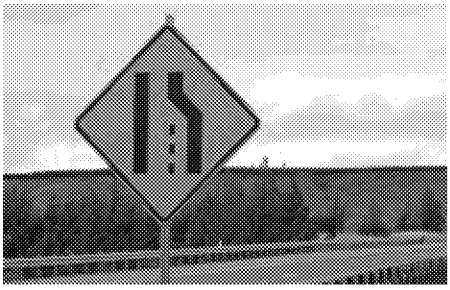
Further Information:

- Federal Register

EPA 'narrowing' scope of first ten TSCA risk evaluations

Agency to exclude uses regulated under other laws, additional consumer exposures

7 June 2018 / TSCA, United States



The US EPA has made it clear it does not intend to evaluate exposure routes it considers adequately regulated under other laws. This narrows the scope of the assessments of the first ten substances subject to risk evaluation under the amended TSCA.

Problem formulation documents, released on 1 June, identify the 'conditions of use' the agency plans to evaluate and serve as an interim step toward final risk evaluations, which must be completed by December 2019.

As was the case in the preliminary 'scoping documents', published last June, the EPA plans to exclude 'legacy' uses of chemicals from its risk evaluations. It has since identified additional uses that are no longer ongoing, which also will be set aside.

Which 'conditions of use' the agency considers in its risk evaluations remains heavily disputed. A coalition of NGOs is suing to challenge the Trump administration's interpretation that it has the discretion not to evaluate all of a substance's conditions of use.

Previously addressed

In its final framework rules, the EPA said it may "exclude a condition of use that has been adequately assessed by another regulatory agency, particularly where the other agency has effectively managed the risks".

The problem formulation documents indicate the agency is interpreting this to cover chemical exposures the EPA itself regulates. For example, the evaluations will not consider the ozone depletion potential of carbon tetrachloride, methylene chloride and 1-bromopropane because these risks are "adequately assessed and effectively managed under the Clean Air Act (CAA)".

Of the ten chemicals being evaluated, only the dye Pigment Violet 29 is not affected by any such exclusions.

More uses trimmed

The problem formulations also reflect several 'conditions of use' that had been included in last year's scoping documents, but have since been removed after the EPA concluded they are no longer active in commerce.

These include consumer uses of several chemicals that the agency determined are no longer an issue or appear only in incidental amounts. Examples of such exclusions include:

- consumer exposures to 1,4-dioxane that were considered "minimal";
- carbon tetrachloride in cleaners and degreasers intended for consumer use; and
- consumer exposures of 1-BP in degreasers for engine care.

'As narrow as they could be'

The approach reflected in the problem formulations makes the TSCA risk evaluations "about as narrow as they could be," Melanie Benesh, legislative attorney for the Environmental Working Group (EWG), told Chemical Watch.

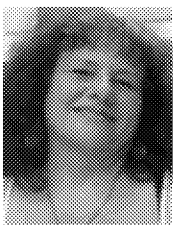
And she questioned some of the decisions regarding consumer products. For example, 1,4 dioxane does appear as an unintentional contaminant, she said, and the EPA was given that information in comments though the problem formulations say otherwise.

"We are very concerned about taking away these really significant sources of exposure" from evaluations, she added.

By not evaluating uses that are otherwise regulated, the "EPA will effectively assume that such exposures to known toxic chemicals pose no risk," Richard Denison, lead senior scientist at the Environmental Defense Fund (EDF), said in a blog post.

This amounts to delegating chemical regulation piecemeal to various EPA offices "with no obligation for those other EPA offices to act," he said.

Comments on the problem formulation documents will be accepted for 45 days following their publication in the *Federal Register*.



Julie Miller

Reporter

Related Articles

- EPA names first ten chemicals for new TSCA evaluations
- EPA issues TSCA 'problem formulation' documents
- US EPA issues final TSCA framework rules
- NGOs, Democrats decry 'weakened' TSCA framework rules
- NGOs focus on 'conditions of use' in TSCA framework rules suit
- US EPA issues final TSCA framework rules

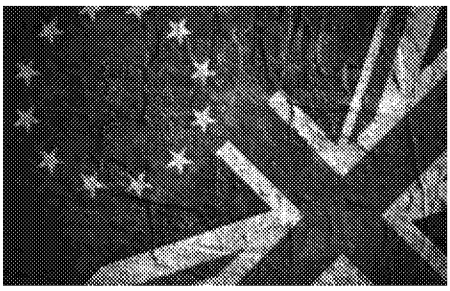
Further Information:

- Problem formulations

Brexit showing 'little impact' on 2018 REACH registrations – Hansen

UK submissions figures 'as expected' for final deadline

7 June 2018 / Europe, REACH, Substance registration, United Kingdom



Registration figures under the 2018 REACH deadline show there is "no indication" that UK companies are not fulfilling their obligations under the Regulation, Echa head Bjorn Hansen said.

After Germany, registrations submitted by UK companies were the second highest with 4,515 dossiers covering 2,020 substances.

These figures, Mr Hansen said during a media briefing last week, are "fully within the range of what we had expected".

The UK voted to leave the EU in June 2016. With an expected withdrawal date of 29 March next year and with no guarantees of remaining within REACH and associated with Echa, UK industry had raised questions about whether or not to register.

One option for those companies to ensure continued access to the single market and ongoing validity of their registrations is to relocate or appoint only representatives in another EU country.

According to a Chemical Watch survey conducted early this year, a third of UK-based companies said they are actively organising or planning to move some of their operations out of the country because of the regulatory uncertainty around Brexit.

However, Mr Hansen said that Echa has not seen "any major effect of any companies moving location".

What the agency has seen, he said, is "significant" use of Echa's Brexit webpage which contains advice and Q&As regarding post-Brexit implementation of REACH in the EU27.

"We've received a lot of questions via our help desks so I would say my feeling is UK companies are fully aware that something could happen and they are definitely informing themselves".

Supply chain

One of the main issues of the UK being outside of REACH, Mr Hansen said, lies with companies and their supply chains between the EU and Britain. Another is for UK companies "who have invested in their registration dossiers and are, of course, expecting to reap in the benefits of the internal market that comes from that investment".

He identified two areas that industry should consider. The first is to check if the registrants are in the UK. The second relates to mixtures being shipped between the EU and the UK. "Looking at the sizes of the different markets, there is clearly an issue for UK companies who want to maintain access to the European market through a registration," he said. "They will need to do some more work in tracking where their chemical is going within the UK and through the EU."

Asked whether it would be possible for the UK to have associate membership of Echa, Mr Hansen said "anything is possible, but it all depends on what our political masters representing the EU27 and those representing the UK agree".

He added that he is "very confident" that Echa will be able to implement whatever they agree. "Depending on the technicalities we might need a little more time, but principally we don't see any big technical or scientific or administrative problems in implementing what our political masters decide is the right relationship with the UK."

Non-EU registrants

Twenty-nine percent of registrations for the 2018 deadline for chemicals of 1-100 tonnes by volume were made by only representatives on behalf of a non-EU companies. And 48% were made by importers – some of which may also be manufacturers. The rest of the registrations were submitted by EU/EEA producers.

Echa's director of registration Christel Musset said that generally for the previous deadlines Echa had "virtually the same number" of registrations for chemicals produced in Europe and chemicals imported into Europe. "But for the last deadline it has been different. We had many more registrations from importers".

The top five importers of REACH registered chemicals for the 2018 deadline are:

- USA 27%;
- China 15%;
- Japan 14%;
- India 12%; and
- Switzerland 6%.

The agency recently conducted a survey looking at the number of substances on the US market, which Echa said is similar to Europe in terms of size and structure. It found the number of substances on that market is "very similar". It was "comforting", Ms Musset said, that companies based there registered their substances with Echa in a timely fashion.

Luke Buxton

Europe desk editor

Related Articles

- ORO publishes new Brexit advice for only representatives
- Brexit uncertainty forcing UK-based firms to act

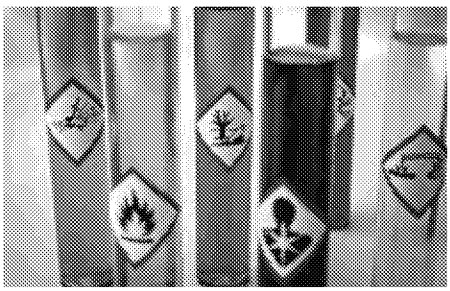
Further Information:

- REACH 2018 results
- Echa REACH 2018 media briefing
- Echa Brexit webpage

EU trade body calls for simplified cleaning products labels

Survey shows decline in consumers reading on-pack information

7 June 2018 / Cleaning products, Europe, Labelling



European soap and detergents trade body, Aise, is calling for cleaning product labels to be simplified, following its consumer behaviour survey which claims fewer people are reading information on packaging.

The survey of more than 4,600 people in 23 countries found that only half read the labels before buying a product and 60% before using it.

This figure has fallen since Aise's first consumer survey in 2008, which found 68% read the label before purchase and 75% before use.

Aise president, Arndt Scheidgen said that overlaps between the detergents Regulation, classification, labelling and packaging (CLP) and biocides Regulations have resulted in "crowded labels".

According to the CLP Regulation, the label of a product containing hazardous chemicals must include information on the supplier, the quantity of substances or mixtures and, where relevant, hazard pictograms, signal words, hazard and precautionary statements.

But Mr Scheidgen said "too much and redundant information is counterproductive to safe use".

In a press release, he calls for the upcoming detergents Regulation review to prioritise "simplification of labels to enable safe and correct use of products" and "the transparent provision of information via digital means". This is due to take place in 2019.

Simplified labels

In 2016 and 2017, Aise carried out research into product labelling, as part of its Better Regulation and Safe Use project.

This concluded that consumers preferred simplified labels and found icons or precautionary pictograms easier to understand than text.

Aise communications director, Valérie Séjourné, told Chemical Watch the association was keen to explore simplified label designs with the Commission and other stakeholders.

"The idea is to be much more efficient, relevant with regards to consumer information, more impactful, and tap into the opportunity of leaving what really matters on the pack, to allow good information on good use, safe storage and allergy information," she said.

Aise has designed a pictogram to replace the written warning 'Keep out of the reach of children', which is under discussion with the UN.

'Unrealistic expectations'

But the European consumer organisation, Beuc, says that full ingredients information and dosage instructions should stay on packaging rather than move online.

Beuc's safety expert, Pelle Moos, told Chemical Watch: "For safety reasons, it is essential that allergen messages stay on the label if we want to protect consumers from skin rashes and other irritations.

"Because consumers have busy lives, it is unrealistic to expect them to search for information on their phone while in the supermarket."

A Commission spokesperson confirmed that a support study was conducted into the detergents Regulation, which concluded in January.

The study will be made public together with a staff working document later this year.

They added that comments and suggestions made by Aise and other consulted stakeholders will be taken into account.

Tammy Lovell

Business reporter

Related Articles

- EU Commission consults on detergents Regulation
- Classification of household detergent products under CLP: challenges and opportunities
- EU chemicals regulations 'failing' consumers, member states say
- EU Commission pressed to review detergents labelling

Further Information:

- Aise press release
- Detergents Regulation (EC) No 648/2004
- Better Regulation and Safe Use project (BRES)

Norway updates product registration tool

7 June 2018 / Norway

Norway's environment agency has updated its electronic tool for declaring chemicals to the product register.

The service was launched for domestic companies in 2015, and opened to foreign companies this year.

The agency says it has made several simplifications, improvements and fixes, partly in response to user feedback.

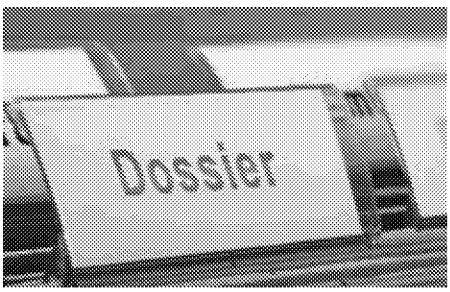
Further Information:

- Press release (in Norwegian)

REACH registered substances total 'worse than expected' – Ueapme

Shortfall in figures an 'indication we have a problem'

7 June 2018 / Europe, REACH, Substance registration



The number of substances registered under all three REACH deadlines is "worse than expected", according to European SMEs trade body Ueapme.

On 1 June – the day after the third and final deadline – Echa confirmed that 21,551 substances have been registered.

This number is short of the estimated total from a 2001 European Commission White Paper of 30,000 substances. Meanwhile, Echa had later predicted that 25,000 new chemicals would be registered for the 1 to 100 tonne deadline alone. The final number was 6,824

Ueapme advisor Marko Sušnik says the gap between the White Paper estimate and the final registered number is considerable and "it is hard to believe that the market has changed so drastically in the last ten to 15 years".

The number of 21,551 substances cannot be "easily compared" to the White Paper's estimation, he says. "First, one needs to deduct most non-phase-in substances and all on-side isolated intermediates. Based on this, it is highly probable

that we do not even have 15,000 relevant substances registered." This would mean, he said, that not even half of the 30,000 estimate would be covered by registrations.

In a bid to quell concerns prior to the final registration deadline, Echa and Cefic told companies there was no need to panic over 'missing' substances because the estimate was made over 15 years ago and reflected the market at the time.

But Mr Sušnik, who has long raised concerns over 'missing' registered REACH chemicals, says a really valuable exercise would be for someone to analyse the numbers.

The lower-than-expected number of chemicals registered under REACH should be considered alongside those on other European inventories to get an idea of what is on the market, he says.

The classification and labelling inventory (C&L), which has been in place since 2011, contains about 135,000 individual substances. And this total is similar to the sum of the substances contained in:

- the European Inventory of Existing Commercial Chemical Substances (Einecs) that were deemed to be on the European market between 1971 and 1981;
- the European List of Notified Chemical Substances (Elincs); and
- the no-longer polymers (NLP) list.

From these, Mr Sušnik says "you get a taste of what is on the EU's market right now and this is not so different from a decade (and more) ago, at least in numbers".

'Problem' for SMEs

According to Echa figures, SMEs made up about 18% of registrants submitting dossiers for all REACH three tonnage bands. For the 2018 deadline for 1 to 100 tonnes, they made up 17%.

However, it was widely expected that SMEs would comprise the largest number of registrants for the final lower volume deadline, Mr Sušnik says. The result is "an indication we have a problem".

It means, he says, "companies that should have registered, have not registered and it becomes worrisome. Now we have a certain instability and no one really understands what is happening.

"Either there is something going on there or we have completely wrong estimations." The latter, he says, would be a "best case" scenario. "If you find out in half a year that the estimation is wrong then fine – but so far there is no convincing argument why it should be wrong because all the other numbers are pretty accurate."

He refers to the fact that estimations of substances to be registered for the 2010 and 2013 were close to the final numbers. This is corroborated by Cefic's REACH director Erwin Annys, who in February said "everyone was astonished" at how accurate the estimation was for both of these deadlines.

Mr Sušnik admits there have been "discrepancies" in the estimates, but "not huge ones. If you were talking about 100 or 200 substances you can say 'forget it, we can handle this', but we are talking about thousands."

Industry stance

Speaking on 1 June, Echa's director of registration Christel Musset said "it's a bit too early to say" if Echa should be concerned about the shortfall because the agency will not have the final registration figures until after the completeness checks are finished in August.

And in comments to Chemical Watch, Cefic said: "Despite the data showing fewer registrations than was initially expected, we are convinced that the absolute majority of the substances currently used on the market have been successfully registered."

Luke Buxton

Europe desk editor

Related Articles

- Industry concerned about 'missing' REACH 2018 substances
- 'Don't panic' – Echa, Cefic address 'missing' REACH substances
- Industry concerned about 'missing' REACH 2018 substances
- No convincing answers on 'missing' REACH 2018 substances – Ueapme

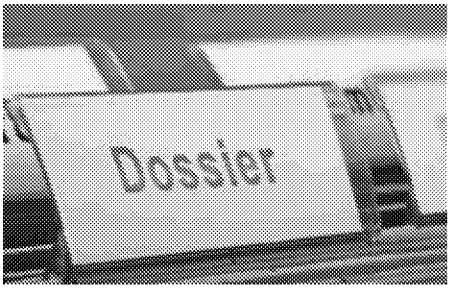
Further Information:

- 2001 European Commission White Paper
- REACH registration results

REACH registered substances total 'worse than expected' – Ueapme

Shortfall in figures an 'indication we have a problem'

7 June 2018 / Europe, REACH, Substance registration



The number of substances registered under all three REACH deadlines is "worse than expected", according to European SMEs trade body Ueapme.

On 1 June – the day after the third and final deadline – Echa confirmed that 21,551 substances have been registered.

This number is short of the estimated total from a 2001 European Commission White Paper of 30,000 substances. Meanwhile, Echa had later predicted that 25,000 new chemicals would be registered for the 1 to 100 tonne deadline alone. The final number was 6,824

Ueapme advisor Marko Sušnik says the gap between the White Paper estimate and the final registered number is considerable and "it is hard to believe that the market has changed so drastically in the last ten to 15 years".

The number of 21,551 substances cannot be "easily compared" to the White Paper's estimation, he says. "First, one needs to deduct most non-phase-in substances and all on-side isolated intermediates. Based on this, it is highly probable that we do not even have 15,000 relevant substances registered." This would mean, he said, that not even half of the 30,000 estimate would be covered by registrations.

In a bid to quell concerns prior to the final registration deadline, Echa and Cefic told companies there was no need to panic over 'missing' substances because the estimate was made over 15 years ago and reflected the market at the time.

But Mr Sušnik, who has long raised concerns over 'missing' registered REACH chemicals, says a really valuable exercise would be for someone to analyse the numbers.

The lower-than-expected number of chemicals registered under REACH should be considered alongside those on other European inventories to get an idea of what is on the market, he says.

The classification and labelling inventory (C&L), which has been in place since 2011, contains about 135,000 individual substances. And this total is similar to the sum of the substances contained in:

- the European Inventory of Existing Commercial Chemical Substances (Einecs) that were deemed to be on the European market between 1971 and 1981;
- the European List of Notified Chemical Substances (Elincs); and
- the no-longer polymers (NLP) list.

From these, Mr Sušnik says "you get a taste of what is on the EU's market right now and this is not so different from a decade (and more) ago, at least in numbers".

'Problem' for SMEs

According to Echa figures, SMEs made up about 18% of registrants submitting dossiers for all REACH three tonnage bands. For the 2018 deadline for 1 to 100 tonnes, they made up 17%.

However, it was widely expected that SMEs would comprise the largest number of registrants for the final lower volume deadline, Mr Sušnik says. The result is "an indication we have a problem".

It means, he says, "companies that should have registered, have not registered and it becomes worrisome. Now we have a certain instability and no one really understands what is happening.

"Either there is something going on there or we have completely wrong estimations." The latter, he says, would be a "best case" scenario. "If you find out in half a year that the estimation is wrong then fine – but so far there is no convincing argument why it should be wrong because all the other numbers are pretty accurate."

He refers to the fact that estimations of substances to be registered for the 2010 and 2013 were close to the final numbers. This is corroborated by Cefic's REACH director Erwin Annys, who in February said "everyone was astonished" at how accurate the estimation was for both of these deadlines.

Mr Sušnik admits there have been "discrepancies" in the estimates, but "not huge ones. If you were talking about 100 or 200 substances you can say 'forget it, we can handle this', but we are talking about thousands."

Industry stance

Speaking on 1 June, Echa's director of registration Christel Musset said "it's a bit too early to say" if Echa should be concerned about the shortfall because the agency will not have the final registration figures until after the completeness checks are finished in August.

And in comments to Chemical Watch, Cefic said: "Despite the data showing fewer registrations than was initially expected, we are convinced that the absolute majority of the substances currently used on the market have been successfully registered."

Luke Buxton

Europe desk editor

Related Articles

- Industry concerned about 'missing' REACH 2018 substances

- 'Don't panic' – Echa, Cefic address 'missing' REACH substances
- Industry concerned about 'missing' REACH 2018 substances
- No convincing answers on 'missing' REACH 2018 substances – Ueapme

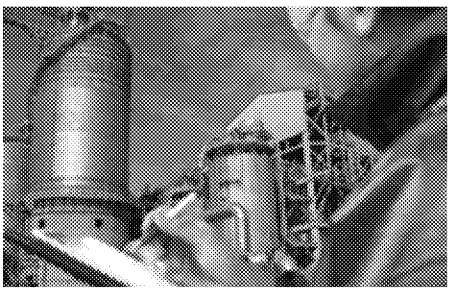
Further Information:

- 2001 European Commission White Paper
- REACH registration results

South Korea releases K-REACH draft implementation rules

Pre-reporting required by 30 June 2019

7 June 2018 / Enforcement, K-REACH, South Korea, Substance registration



South Korea's environment ministry (MOE) has published draft implementation rules for the updated K-REACH legislation.

The rules, released on 30 May, include details on: pre-reporting substances; tonnage bands registration dates; simplified registration; and confidential business information (CBI) exemptions for manufacturers/importers.

Consultation on them closes on 9 July.

Pre-reporting

According to the rules, companies are obliged to 'pre-report' the substances they manufacture or import by 30 June 2019. This will include information on classification and labelling and anticipated volumes. If a company does not do this it will, in effect, be banned from manufacturing or importing from 1 July 2019.

Pre-reporting a substance allows a company to register to the given deadline – what the MOE calls a grace period.

Pre-reporting will take place through an online system to the Korea Environment Corporation (Keco) – an MOE agency. The ministry will publish detailed rules for this at a later date.

Tonnage band deadlines

The first registration deadline of 31 December 2021 is for:

- substances manufactured/imported in total volumes of 1,000 tonnes or more a year; and

- high concern substances, for example carcinogenic, mutagenic or reprotoxic (CMR), made or imported in total volumes of one tonne or above a year.

These substances will be designated and notified by the MOE by the end of the year.

However, the rules stipulate that substances below one tonne can also be subject to registration according to MOE hazard and risk assessment.

Simplified registration

In general, substances can qualify for simplified registration if:

- low concern substances – those not classified as hazardous for human or environmental health under the UN GHS; and
- intermediate substances that are completely used up in production.

For low concern substances, 15 hazardous test data endpoints are stipulated for simplified registration – in contrast to the 47 required for full registration.

CBI exceptions for downstream information

The rules also amend confidential business information (CBI) exemptions for supplying information to downstream users.

Currently manufacturers and importers must supply information to purchasers of hazardous substances by default. However, it is possible to gain an exemption from the MOE by "pre-approval", they say.

They also extend the information requirement to CMR substances, where the name, contents, hazard classification, and labelling/marks must be supplied.

More details available on CW+AsiaHub

Sunny Lee

Asia editor

Related Articles

- Updated K-REACH set for 20 March adoption
- Increased penalties as K-REACH amendment passes parliamentary committee
- South Korea's draft implementation rules arrive for updated K-REACH

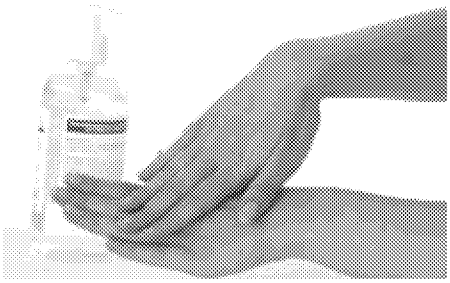
Further Information:

- MOE announcement (in Korean)

Triclosan may worsen gut inflammation and trigger cancer, study suggests

Chemical may alter populations of gut bacteria

7 June 2018 / Biocides, China, Risk assessment, United States



The antimicrobial chemical triclosan may promote inflammation in the colon and trigger associated colon cancer, according to a rodent study. The inflammation appears to result from triclosan's effects on bacterial populations in the gut.

A team from the US and China found that triclosan – at concentrations reported in human blood – caused a low level of gut inflammation in healthy mice. They then used chemicals to induce gut inflammation before feeding mice food laced with triclosan. They also engineered some mice to develop inflammatory bowel disease while in others they administered chemicals known to cause colon cancer.

Triclosan increased the severity of colitis symptoms and "spurred" colitis-associated colon cancer cell growth, according to the report in *Science Translational Medicine*.

Led by Haixia Yang and Weicang Wang at the University of Massachusetts, the researchers also carried out tests on germ-free mice – which have no gut bacteria – finding that triclosan had no inflammatory effect. This reinforces the idea that possible harmful effects may result from changes to gut bacteria, they suggest.

Further studies are "urgently needed", write the researchers, who are also concerned about other, less documented, anti-microbial compounds.

In 2016, more than 200 scientists and medical professionals published the so-called Florence Statement, calling for governments and industry to limit the use of triclosan and triclocarbon in consumer products. The chemicals should only be used when they provide an evidence-based health benefit, said the statement. The signatories pointed to evidence that the chemicals persist in the environment and can harm aquatic organisms.

In Europe, triclosan has been banned from use in human hygiene products since 2017. The European Commission has also amended the EU cosmetics Regulation to restrict triclosan to a maximum concentration of 0.3% in products, including toothpastes, hand soaps and body soaps. It is permitted in mouthwashes at a maximum concentration of 0.2%.

The US FDA recently banned triclosan from use in over-the-counter healthcare antiseptics.

Related Articles

- Scientists issue global call to act on triclosan, triclocarban

- EU bans five parabens, restricts triclosan in cosmetics

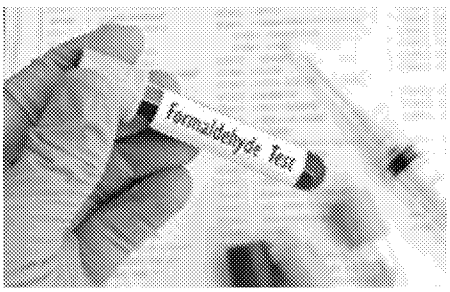
Further Information:

- Journal article

ACC fears 'restructuring' of IRIS formaldehyde assessment

Long-awaited final version must 'revisit science', it says

7 June 2018 / CMRs, Risk assessment, Toxicology, United States



The American Chemistry Council has written to the EPA voicing concern over the long awaited IRIS programme's chemical assessment of formaldehyde. In the letter sent earlier this year, but shared with *Chemical Watch* last week, the ACC said it feared the final assessment would merely be a "restructuring" of the original draft and not a revisiting of the science.

The final version has been in the pipeline for eight years, after the draft received criticism from the National Academies of Sciences, Engineering, and Medicine (NAS) and industry. Last month, three Senate Democrats suggested in a letter to the EPA that the programme had actually completed its "revised draft" last autumn, but was stalling its release under pressure from industry.

The ACC says the programme must revisit the science, including consideration of mode of action, to ensure it overcomes the shortcomings of the draft from 2010.

But, during a meeting in January, National Center for Environmental Assessment (NCEA) and IRIS staff allegedly told the ACC they were not doing this. The ACC then wrote to the EPA on 26 January.

Revisiting the science

The science in question concerns leukaemia, and a suggested causal link with formaldehyde exposure. In its draft assessment, the IRIS programme concluded that there was evidence of this.

However, this conclusion was disputed by industry. And in its 2011 review of the assessment, the NAS was highly critical of the process that was used.

The IRIS programme has since embarked on a redrafting exercise, but there is disagreement about to what extent it needs to revisit the science.

Jennifer Sass at the Natural Resources Defense Council (NRDC), a US NGO, says that the NAS did not question the conclusions of the assessment and that this is key. The NAS was concerned first and foremost that the IRIS programme had not shown exactly how it arrived at its conclusion. "Essentially, they didn't show their work," Dr Sass says.

But the ACC says that in the 2010 draft the IRIS programme failed to properly integrate all the evidence streams, and this failure – which the organisation says was clearly identified by the NAS – led to a false conclusion: that formaldehyde causes leukaemia.

Three data streams

The scientific evidence can be divided into three data streams:

- epidemiological (human);
- animal; and
- mode of action.

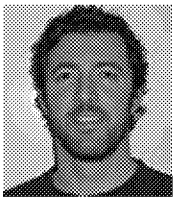
The human data – derived from studies of workers exposed to formaldehyde – is, at best, inconsistent, with some studies showing no link and others showing a weak link, says Kimberly White, who manages the ACC's formaldehyde panel. Furthermore, those that do show a link, such as the Zhang study, are flawed.

Bernard Goldstein, a physician and emeritus professor of environmental and occupational health at the University of Pittsburgh, says he remains "less than convinced" that formaldehyde causes leukaemia. But he is nonetheless critical of many aspects of the ACC position.

Professor Goldstein was part of the International Agency for Research on Cancer (Iarc) working group that assessed formaldehyde in 2012. The group was split on the leukaemia risk. The majority found the evidence "sufficient" and a minority, which included Professor Goldstein, found it "limited".

But he says industry has taken the wrong approach to the human data. Instead of reanalysing the data from the Zhang study, the study should have been repeated with a different cohort to eliminate potential confounding factors.

Industry funded reanalysis of the data, published last year, found "no association between individual formaldehyde exposure estimates and frequency of aneuploidy [which is indicative of leukaemia] among those exposed". But Professor Goldstein criticised the "Mundt paper" in a letter to the editor, in which he said that, if anything, it reinforced the original findings.



Andrew Turley

Science editor, Chemical Watch

Related Articles

- Industry accused of stalling US EPA formaldehyde assessment
- ACC calls for inclusion of industry analysis in formaldehyde assessment

Further Information:

- ACC letter

- EPA reply
- Professor Goldstein's letter about Mundt paper

Dollar General CEO to meet with chemical campaigners

'Step in the right direction' for discount retailer

7 June 2018 / Bisphenols, PFCs, Phthalates, Retail, United States, Voluntary action

Todd Vasos, the CEO of US discount chain Dollar General, has agreed to meet with campaigners to discuss their concerns about hazardous chemicals in the stores' products.

Activists from the Campaign for Healthier Solutions (CHS) network of environmental justice groups, bought shares in Dollar General so they could attend its annual shareholders' meeting on 31 May.

They used the meeting as an opportunity to question executives about chemicals in the company's products and its lack of a chemicals policy.

Although they did not receive an immediate answer to their concerns, Mr Vasos agreed to meet with campaign representatives at a future date.

Eric Whalen of CHS told Chemical Watch: "Dollar General, one of the nation's leading retailers, is one of the last remaining stores that hasn't taken action on chemicals yet, so the appearance that they're starting to make moves in the right direction is definitely good."

CHS campaign coordinator Jose Bravo, said that until Dollar General announced a publicly accountable chemical policy, it will continue to fall further behind other stores and "increasingly risk becoming a liability to its shareholders".

Disproportionate impact

CHS focuses on campaigning to reduce chemical exposures from the growing retail sector of dollar stores, which have lagged behind other major retailers such as Walmart and Target in implementing chemical policies.

Discount stores were rated among the lowest performers in last year's retailer 'report card' from the Mind the Store NGO campaign, which evaluated chemical policies from 30 US retailers.

Dollar General ranked bottom with an F rating, while discount retail chain Dollar Tree ranked 16th overall, with a D score.

Mind the Store campaign director, Mike Schade, said that although meeting with campaigners was a step in the right direction for Dollar General, it was "no substitute for effective public chemical safety policies."

The agreed meeting is the latest victory for the CHS campaign after fellow discount store Dollar Tree committed to the elimination of 17 chemicals of high concern from its products by 2020.

Tammy Lovell

Business reporter

Related Articles

- US NGOs pressurise discount retailers over PFAS in popcorn bags
- Walmart aligns disclosure policy with Californian law
- Target takes aim at 'unwanted chemicals'
- Apple comes top in US retailer chemical ranking
- Dollar Tree criticised for lack of action on chemicals

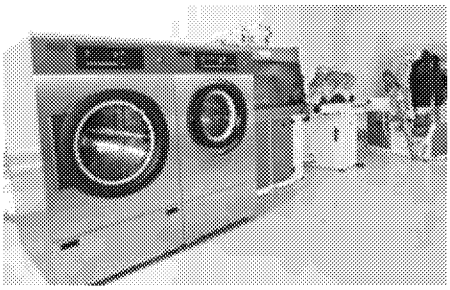
Further Information:

- CHS press release
- CHS report
- Dollar Tree letter

US 'problem formulations' raise fears for TCE, NMP rules

EPA to reevaluate previously assessed uses

7 June 2018 / Solvents, TSCA, United States



Concerns have been raised that recently released 'problem formulation' documents confirm the US EPA will abandon pending rules banning some uses of trichloroethylene (TCE) and N-methylpyrrolidone (NMP).

The new documents, published by the agency on 1 June, clarify how the EPA will conduct risk evaluations on the first ten chemicals to be assessed under the amended TSCA.

But they also have implications for proposed rules on three of those substances – TCE, NMP and methylene chloride. Published in the final days of the Obama administration, two of these seek to prohibit the use of TCE in vapour degreasing, and as an aerosol degreaser and spot cleaner in dry cleaning. Another would ban the use of NMP and methylene chloride in consumer and commercial paint strippers.

After a period in which these proposals seemed to be dropping off the agenda, the rule on methylene chloride now seems certain to progress. Its changed trajectory follows a May meeting between EPA Administrator Scott Pruitt and families of people who died as a result of using paint strippers containing the substance.

However, Melanie Benesh, legislative attorney for NGO the Environmental Working Group, told Chemical Watch the problem formulations suggest that the pending bans for uses of TCE and NMP will not be similarly revived.

"It seems pretty clear those rules are not going to be finalised," she said.

Changed course

The changed tack comes after preliminary 'scoping documents', published in June 2017, said the EPA would not carry out further analysis of the uses covered by the proposed rules as it awaited the pending regulations.

But the latest documents indicate that the EPA's upcoming risk evaluations of TCE and NMP will take a fresh look at those previously evaluated uses, to make the assessments "more robust".

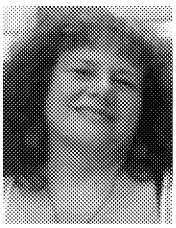
"This means that it could be years before the EPA prohibits these uses," said Ms Benesh. Or the agency "could also scrap the bans altogether".

The problem formulation for methylene chloride does not similarly call for a reevaluation of its paint stripping use. It instead cites plans to address that use by finalising its proposed risk management rule.

"They were on the same path with methylene chloride" until EPA Administrator Scott Pruitt met with the families of victims, said Richard Denison, lead senior scientist at the Environmental Defense Fund. He "apparently changed course," on this substance.

But when the EPA signalled plans to proceed on methylene chloride, it did not indicate whether the forthcoming action would also address NMP – a replacement solvent that was also covered in the original proposal.

The divergent treatment of paint stripping in the two substances' problem formulations suggests they are likely to be considered separately.



Julie Miller

Reporter

Related Articles

- EPA issues TSCA 'problem formulation' documents
- Campaigners push US EPA to finalise TCE ban
- EPA names first ten chemicals for new TSCA evaluations
- US EPA moves to ban additional use of TCE
- US EPA proposes first substance ban in 27 years
- US EPA proposes prohibitions on methylene chloride, NMP
- US EPA commits to act on methylene chloride paint strippers
- US EPA issues final TSCA framework rules

Further Information:

- Problem formulations
- EPA response to comments

Echa round-up

7 June 2018 / Alternatives assessment & substitution, Europe, REACH

New intention to identify an SVHC

Echa has received an intention from Sweden to identify 2,2-bis (4'-hydroxyphenyl)-4-methylpentane as a substance of very high concern due to its suspected reprotoxic properties. The proposal is expected to be submitted by 6 August.

Help with submission difficulties

The agency has reminded registrants of ways to access help with REACH registration problems. They are:

- REACH-IT is open every day, 24 hours a day;
- REACH 2018 Support live chat is open on weekdays from 9:00 to 16:00 CEST until 8 June; and
- questions can also be posed via the agency's contact form.

Video: last-minute advice on REACH 2018 registration

The agency has released the video of its last-minute advice Q&A session on REACH 2018 registration. Questions were posed to an agency expert panel.

Further Information:

- Registry of SVHC intentions
- Video on last minute advice on REACH 2018 registration

New York finalises cleaning products disclosure policy

Intentionally added ingredients must be disclosed by July 2019

7 June 2018 / Cleaning products, Confidentiality & right-to-know, United States

New York's Department of Environmental Conservation has published its finalised cleaning product disclosure policy, which sets out information requirements for manufacturers.

Manufacturers will be required to post specific information about their products and the ingredients they contain on their websites.

Adoption of the policy - the Household Cleansing Product Information Disclosure Programme - follows an announcement last year that the state would begin to mandate ingredient disclosure from the sector. It covers soaps and detergents, containing a surfactant as a "wetting or dirt emulsifying agent and used primarily for domestic or commercial cleaning purposes, including but not limited to the cleansing of fabrics, dishes, food utensils and household and commercial premises".

DEC Commissioner Basil Seggos said the policy will "help the state better understand what chemical hazards the public is exposed to, especially from products made in countries with less protective environmental laws than the US, and reduce exposure to chemicals of concern".

Ingredients

The manufacturer's webpage must list intentionally added and nonfunctional ingredients in the product, by weight in a descending order.

The function of intentionally added ingredients must also be provided, such as stating it acts as a surfactant, colourant, preservative or fragrance. Non intentionally added ingredients should be reported as non functional ingredients, byproducts or contaminants.

Manufacturers must also identify the presence of any 'chemicals of concern'. These include substances found on a number of 'authoritative lists' set out in the policy.

These include:

- those identified by the US EPA as chemicals of concern, PBTs, priority chemicals, or ozone-depleting substances;
- the EU Endocrine Disruptor list and list of substances of very high concern (SVHCs);
- substances listed as reproductive toxicants or carcinogens under California's Proposition 65;

- chemicals classified as a group 1, 2a or 2b by the International Agency for Research on Cancer (Iarc); and
- those ingredients classified under the GHS as skin or eye irritants, respiratory or skin sensitisers, mutagens or aquatic toxicants.

An ingredient in a product must be disclosed if it is present on one or more of the authoritative lists of chemicals of concern. This includes those being withheld as confidential business information (CBI).

Manufacturers are also required to list any nanoscale materials, as defined under the new TSCA. This includes any substance that is processed to form particles in the size range of 1-100 nanometers.

They must also post on their websites any investigations or research they have carried out or commissioned on the effects the products or ingredients have on human health and the environment.

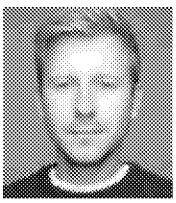
Deadlines

Ingredients intentionally added to these products, present above trace quantities, must be posted online by 1 July 2019. This does not apply to fragrance and nonfunctional ingredients – such as byproducts and contaminants – which must be disclosed by 1 July 2020.

However, manufacturers with 100 employees or fewer have until the later July 2020 deadline to make available information on intentionally added ingredients.

The DEC said in a press release that the policy - borne from a forgotten and unenforced 1970s law - is expected to "serve as an example that can be expanded into other sectors of public disclosure or mirrored by other governments".

Ingredient disclosure has been a focus for both state and federal legislators in recent years. And fragrance ingredients have been a key industry issue, with the likes of Unilever, Procter & Gamble and SC johnson all increasing disclosure efforts.



Leigh Stringer

Global Business Editor

Related Articles

- New York to collect data on cleaning product ingredients
- US House bill introduced on cleaning product ingredients
- Industry groups prepare for ingredient disclosure battle
- Unilever US to disclose fragrance ingredients to consumers
- Procter & Gamble to reveal all fragrance ingredients by 2020
- SC Johnson begins disclosing fragrance ingredients

Further Information:

- Policy
- DEC press release

© 2018. Reprinted and distributed by kind permission of Chemical Watch.

OTHER ARTICLES

Lowe's Listens, Drops Dangerous Toxics; Home Depot Hedges

Natural Resources Defense Council

Scott Pruitt and the Trump Administration's strategy of stalling and stonewalling on action to protect the public from toxic chemicals is starting to ...

Consumer Product Safety Testing Overlooks Cancer Risk From Exposure to Multiple Chemicals

EcoWatch

Exposure to mixtures of **toxic chemicals** often starts in the womb, and the developing fetus gets its first dose of chemical exposure before birth. BPA ...

Lowe's Sends A Strong Message On Hazardous Chemicals In Retail Products

Triple Pundit (blog)

US retailers are beginning to provide consumers with non-toxic alternatives to common household products that contain **hazardous chemicals**. Lowe's ...

September 19, 2017

Comments from Academics, Scientists and Clinicians on the Risk Evaluation Scoping Efforts Under TSCA for Ten Chemical Substances

Submitted online via *Regulations.gov* to dockets: EPA-HQ-OPPT-2016-0736, EPA-HQ-OPPT-2016-0741, EPA-HQ-OPPT-2016-0733, EPA-HQ-OPPT-2016-0723, EPA-HQ-OPPT-2016-0735, EPA-HQ-OPPT-2016-0742, EPA-HQ-OPPT-2016-0743, EPA-HQ-OPPT-2016-0732, EPA-HQ-OPPT-2016-0725, EPA-HQ-OPPT-2016-0737

These comments are submitted on behalf of the undersigned academic, scientists, and clinicians. We declare collectively that we have no direct or indirect financial or fiduciary interest in any chemical under consideration in these risk evaluations. The co-signers' institutional affiliations are included for identification purposes only and do not necessarily imply any institutional endorsement or support, unless indicated otherwise.

We appreciate the opportunity to provide written comments on the scope of risk evaluations for the first ten chemical substances for risk evaluations pursuant to the Toxic Substances Control Act (TSCA), as amended by the Frank R. Lautenberg Chemical Safety of the 21st Century Act (Lautenberg TSCA). Collectively, these chemicals represent an aggregate production volume of more than 1 billion pounds a year in 2015.[1] Some of these chemicals have assessments, and in some cases even restrictions, under other federal programs – but none of these other programs has the mandate given to EPA under the new TSCA: to comprehensively evaluate chemicals and ensure that they do not pose an unreasonable risk to human health and the environment, with special consideration to those most vulnerable amongst us. Therefore, the task ahead for EPA is critical.

These first ten evaluations are also consequential because they will be precedent setting for the implementation of evaluation of science under TSCA. The consequent health impacts of EPA's decisions – for better or worse – will be borne by generations of American children, workers, families, and communities. With so much at stake, we welcome EPA's engagement with the public in this process and we offer EPA concrete approaches to embed the most current scientific principles in its methods to assess the hazards and risks of environmental chemicals.

Our comments address the following main points:

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**
2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**
3. **EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations**

[1] This is the aggregate production volume estimate for the 8 chemicals with production volume information available. For asbestos and pigment violet 29, manufacturers/ importers claimed production volumes as confidential business information (CBI).

ED_004886_00001345-00001

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**
5. **For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.**
6. **Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.**

We are appreciative of the opportunity to provide public input and we look forward to continuing to participate in such opportunities in the near future. Please do not hesitate to contact us with any questions regarding these comments.

Sincerely,

Veena Singla, PhD; Juleen Lam, PhD; and Tracey Woodruff, PhD
Program on Reproductive Health and the Environment
University of California, San Francisco

Patricia D. Koman, MPP, PhD
President and Senior Health Scientist
Green Barn Research

Kathy Attar, MPH
Toxics Program Manager
Physicians for Social Responsibility

Sheila Brear, BDS
Associate Dean of Academic Affairs, School of Dentistry
University of California, San Francisco

Phil Brown, PhD
University Distinguished Professor of Sociology and Health Sciences
Northeastern University

Adelita G. Cantu, PhD, RN
Associate Professor
Alliance of Nurses for Healthy Environments

Courtney Carignan, PhD
Assistant Professor
Department of Food Science and Human Nutrition and Department of Pharmacology and Toxicology
Michigan State University

ED_004886_00001345-00002

Jeanne A. Conry, MD, PhD
President, Environmental Health Leadership Foundation
Past President, The American Congress of Obstetricians and Gynecologists

Carl F. Cranor, PhD
Distinguished Professor of Philosophy
Faculty Member, Environmental Toxicology Graduate Program
Department of Philosophy
University of California, Riverside

Adam M. Finkel, ScD, CIH
Clinical Professor of Environmental Health Sciences, University of Michigan School of Public Health
Executive Director, Penn Program on Regulation, University of Pennsylvania Law School

Susan J. Fisher, PhD
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Robert M. Gould, MD
Associate Adjunct Professor
Program on Reproductive Health and the Environment
University of California, San Francisco
Past President, Physicians for Social Responsibility

Alycia Halladay, PhD
Chief Science Officer
Autism Science Foundation

Jean-Marie Kauth, PhD
Associate Professor
Benedictine University

Detlef R.U. Knappe, PhD
Professor
Department of Civil, Construction, and Environmental Engineering
North Carolina State University

Erica Koustas, PhD
Scientific Consultant
Program on Reproductive Health and the Environment
University of California, San Francisco

Hal C. Lawrence, III, MD, FACOG*
Executive Vice President and CEO
American Congress of Obstetricians and Gynecologists

ED_004886_00001345-00003

Heather Patisaul, PhD
Professor, Biological Sciences
Center for Human Health and the Environment, WM Keck Center for Behavioral Biology
North Carolina State University

Joshua F. Robinson, PhD
Assistant Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Ted Schettler MD, MPH
Science Director
Science and Environmental Health Network

Patrice Sutton, MPH
Research Scientist
Program on Reproductive Health and the Environment
University of California, San Francisco

Aolin Wang, PhD
Postdoctoral Scholar
Program on Reproductive Health and the Environment
University of California, San Francisco

Nsedu Obot Witherspoon, MPH
Executive Director
Children's Environmental Health Network

Lauren Zajac, MD, MPH
Assistant Professor
Department of Environmental Medicine and Public Health
Icahn School of Medicine at Mount Sinai

Marya G. Zlatnik, MD, MMS
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

*indicates organizational support

ED_004886_00001345-00004

DETAILED COMMENTS

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**

Overall, we strongly commend the EPA for its efforts to utilize a systematic and transparent method of research synthesis to reach a concise, strength of evidence conclusion about the human health hazard resulting from exposures to these ten chemicals. Efforts to integrate systematic review methods, including the explicit development of search terms, strategies, and inclusion/exclusion criteria beforehand, is relatively new in EPA's chemical assessment and as such, we applaud the EPA for this and its general improvements in its hazard assessment methodology. These scoping documents generally provide an important infrastructure for outlining EPA's screening approach for identifying relevant references and to document decisions made in the process of identifying the body of scientific literature that will be evaluated in the chemical assessments.

To improve on this document and advance EPA's uptake of systematic review methods of research synthesis, we identify the following opportunities for improvement:

EPA should not exclude studies based on language. EPA's search strategy is limited to English-only studies. The exclusive reliance on English-language studies may not represent the entire body of available evidence, and studies have suggested that language bias might lead to erroneous conclusions.[2] Furthermore, when considering the inclusion or update of an existing systematic review, studies have found that language-inclusive systematic reviews (including studies in languages other than English) were of the highest quality, compared with other types of reviews.[3] Online translation tools are readily available to allow screeners to quickly evaluate study abstracts for relevance, and therefore we recommend EPA to incorporate non-English language studies in their screening and not simply exclude these potentially relevant papers.

EPA should provide exclusion reasons for off topic citations. In the Bibliography Supplemental File for the Scope Documents, EPA has provided lists of bibliographic citations that were identified and screened from the initial literature search and the initial categorization of whether citations were *on topic* or *off topic.* We recommend EPA additionally provide exclusion reasons that were used to come to the conclusion that each citation was *off topic*, as this is a standard recommendation to fulfill transparency in documenting and reporting all decisions made in the study selection process.[4] This is particularly important as EPA has proposed to do its screening in Distiller, proprietary software that presumably will

---

[2] Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., Mierzwinski-Urban, M., Clifford, T., Hutton, B. and Rabb, D., 2012. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *International journal of technology assessment in health care*, *28*(2), pp.138-144.

[3] Moher, D., Pham, B., Lawson, M.L. and Klassen, T.P., 2003. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess*, *7*(41), pp.1-90.

[4] McDonagh M et a. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Agency for Healthcare Research and Quality. 2008.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J. and Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, *6*(7), p.e1000100.

Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. March 2011. Available from: http://handbook-5-1.cochrane.org/chapter_7/7_2_5_selecting_excluded_studies.htm

ED_004886_00001345-00005

not be made publicly available, raising concerns regarding the transparency and reproducibility of this screening step.

EPA should consider other tools for systematic review. EPA has also proposed to extract data results in the DRAGON software. We strongly encourage EPA to also consider other potential software tools that have been developed and actively incorporated into the process of systematic review, such as Swift Reviewer,[5] Active Screener,[6] HAWC (Health Assessment Workplace Collaborative).[7] These tools will help to ensure consistent and transparent execution and presentation of reviews and increase transparency of EPA assessment. Furthermore, we urge EPA to work with the National Toxicology Program and other organizations involved in these efforts in an ongoing basis to develop these and other open source tools to train scientists in their use. We believe that such infrastructure development will be critical to increasing the efficiency of chemical assessments and to expedite uptake of systematic reviews in environmental health.

EPA should have two independent reviewers for screening steps. EPA has outlined its process for screening title and abstracts of papers as having a single reviewer reviewing papers to determine whether the study is *on-topic* or *off-topic*. As part of this process, a senior-level technical expert in the topic area of interest reviewed the appropriateness of the assigned tags for "the first batch of studies" and provided feedback to the screener. Senior-level technicians also provided feedback and guidance on specific references to the individual screeners as needed during the screening and tagging process. From the description of this process, it does not appear that two independent reviewers screen all titles and abstracts for potential inclusion. Using two independent reviewers is a standard approach in systematic reviews and therefore we strongly recommend that EPA include a second independent reviewer within this process to ensure that **all** studies are screened by two reviewers at each step (title and abstract and full text).

EPA should clearly document decisions related to the identification and search. For example, it was unclear how many studies were included in the first batch of studies reviewed by the senior-level technician—these decisions should be clearly specified beforehand as to the number (or percent) that will be reviewed by this independent reviewer. Furthermore, it is unclear how many studies the senior-level technical experts are reviewing generally as to their additional feedback and guidance to individual screeners. This should be more clearly stated and described beforehand in these protocols.
We recommend EPA broaden the set of studies that are initially screened in the first batch to ensure consistency across reviewers and demonstrated understanding of protocol instructions by all reviewers before moving on to screening the remaining records. It is stated in the Gray Literature Search Results that individual screeners would screen and tag 10 references that would be then independently reviewed by the senior-level technical expert. However, this does not seem to be an adequate number of studies as it is a small number relative to the expected number of records that will ultimately be screened.

EPA should clarify how it will handle discrepancies in the inclusion/exclusion and tagging process. As it is stated in its current protocol, it appears that the senior-level technical expert has the final say in

---

[5] Howard, B.E., Phillips, J., Miller, K., Tandon, A., Mav, D., Shah, M.R., Holmgren, S., Pelch, K.E., Walker, V., Rooney, A.A. and Macleod, M., 2016. SWIFT-Review: a text-mining workbench for systematic review. *Systematic reviews*, *5*(1), p.87.
[6] https://www.sciome.com/swift-activescreener/
[7] https://hawcproject.org/about/

determining the final inclusion/exclusion decision and tagging, for the subset of studies they evaluate. However, this should be clarified and we also highly recommend that a third party reviewer be incorporated as an arbiter for these decisions if consensus cannot be reached between the two reviewers, as is typically standard in systematic reviews.

EPA should stratify its exclusion criteria separately at the title and abstract and full-text screening steps. It is likely that title and abstracts of papers would not contain sufficient detail to evaluate all exclusion criteria—many of these would likely only be identified in the full-text of the paper. To increase the efficiency of the screening process, it would help to create a subset of exclusion criteria most relevant when screening the title and abstracts of records versus the larger set of exclusion criteria relevant to screening the full text of records.

EPA should clearly outline the process for handling anticipated overlap with literature relevant to multiple topics. EPA should describe how this will be addressed by the screeners and whether the same reviewer will be responsible for screening papers with inclusion/exclusion criteria across multiple topics or whether different reviewers are responsible only for screening studies for one particular topic. Additional details in regards to the process by which this screening will occur would be helpful. Given the breadth of each assessment (searching literature related to fate, engineering, exposure, human health, and environmental hazard) and the complexity of the screening process (tagging *on-topic* and *off-topic* literature and using additional sub-categories or sub-tags to allow for additional categorization), there appears to be the potential for individual papers to fall into different topic categories and have many different tags and sub-tags applied to indicate their relevance. However, it is unclear how this will be organized in the screening phase. Search strategies and inclusion/exclusion criteria appear to have been developed specifically for each literature topic and the potential overlap of literature relevant to multiple topics is not addressed.

EPA should explicitly include stopping rules for when the list of relevant studies will be considered final. There is no discussion of stopping dates or the process of updating the literature search to search for newer studies. Newer scientific studies will inevitably continue to appear in scientific journals and it will be impossible to continually attempt to include all these studies in a chemical assessment. To meet the deadlines as mandated by the Lautenberg Amendments, EPA should state clear stopping rules in the form of deadlines or criteria for when the body of included relevant studies will be finalized for the purposes of the chemicals assessment.

EPA should ensure gray literature search results are adequately screened. EPA's gray literature search strategy proposes to utilize Google's API to develop custom searches and return the first 100 results, sorted by predicted relevancy so that the results likely to be most relevant are screened first. It is unclear why this number is limited to only the first 100 and whether there was an empirical reason for why this particular number was selected. We recommend that EPA ensure that an adequate number of search results are screened, in particular considering that the gray literature can contribute potentially important information relevant to toxicity, mode of action, exposure, fate and transport, engineering or occupational exposure, or existence of publication bias.

EPA should consider "snowball searching," where the citations of included (i.e., *on-topic*) references are searched as well as using databases such as Web of Science to search for references that cite the included citations. EPA states that it plans on assessing the specificity and efficiency of the literature searches, through comparison of references either cited in existing problem formulation and risk assessment documents, in the public use documents and supporting life cycle diagrams, and comparison

of the references cited in review articles. Snowball searching will contribute to the evaluation of the specificity and efficiency of its literature searches, and also help to identify newer relevant studies that could potentially be included that have not yet been indexed in main databases such as PubMed.

EPA should incorporate appropriate tools for updating and evaluating systematic reviews in their chemical assessments. Garner et al., as part of efforts by a Cochrane Collaboration panel for updating guidance for systematic reviews, published guidance in 2016 for determining when it is appropriate to update a systematic review and outlining the steps for performing the update.[8] We have attached this guidance as an Appendix to these comments. EPA should evaluate the Cochrane tool's applicability to environmental chemicals given that Cochrane systematic reviews are geared towards reviews of clinical intervention evidence, so these tools may require updating and tailoring for an application to environmental health data.

A recent NAS report[9] recommends EPA should develop policies and procedures to allow the agency to identify, use and update existing systematic reviews. The committee also noted that it was important that the existing systematic review's study question directly addresses EPA's topic of interest and that the methods are critically evaluated before the systematic review is used and updated. EPA should ensure that only the highest quality systematic reviews be considered appropriate for use. It will be critical for EPA to develop tools to assist with the process of evaluating existing systematic reviews, particularly as this field continues to rapidly expand and more systematic reviews relevant to environmental health questions are published in the scientific literature, potentially of variable quality.

One tool which might be helpful for evaluating the risk of bias in systematic reviews is the ROBIS tool, which the NAS committee utilized in their report.[10] This tool was developed using rigorous methodology and can be applied for evaluating internal validity of systematic reviews in conjunction with other available tools to critically appraise and assess their quality. Of particular note is the strong emphasis on the recommendation that tools such as ROBIS should not be used to generate a composite quality score, as it has been well-documented that scoring can lead to bias in evaluation of the studies.[11] As such, the ROBIS tool presents several options for visually and graphically presenting results from risk of bias assessments based on individual domains or the overall rating, enabling reviewers to highlight particular areas of concern or reviews that are most relevant to the target question of interest.

Another tool which may be helpful in this process is the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), used by authors of systematic reviews to improve the reporting

---

[8] Garner P, Hopewell S, Chandler J, MacLehose H, Akl EA, Beyene J, Chang S, Churchill R, Dearness K, Guyatt G, Lefebvre C. When and how to update systematic reviews: consensus and checklist. bmj. 2016 Jul 20;354:i3507.

[9] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[10] Id.

[11] Jüni, P., Witschi, A., Bloch, R. and Egger, M., 1999. The hazards of scoring the quality of clinical trials for meta-analysis. *Jama*, *282*(11), pp.1054-1060.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

ED_004886_00001345-00008

of elements relevant to the systematic review and meta-analyses.[12] Increasingly, scientific journals are requiring the inclusion of checklists such as PRISMA with the submission of systematic reviews considered for publication. Although this tool is used to evaluate study reporting, and is not an assessment instrument to gauge the quality of a systematic review, it can still provide a useful framework to identify reported components of an existing systematic review in the process of evaluating quality or to identify missing components requiring follow-up with study authors to obtain additional information.

Furthermore, we strongly encourage EPA to evaluate the potential for financial conflicts of interest as an element in their study design. This is currently included as a consideration in evaluation risks of bias in some frameworks, such as the Navigation Guide,[13] and extracted for consideration as an additional domain in other frameworks, such as that developed by the National Toxicology Program's (NTP) Office of Health Assessment and Translation (OHAT).[14] The Cochrane Collaboration's risk of bias tool does not currently include a specific domain for bias related to study funding source, but this is an area of active discussion among its members.[15] The Cochrane Collaboration has recognized the importance of identifying study funding source, which has been empirically shown to be associated with biases.[16] A recent report from the NRC recommended that the U.S. EPA consider funding sources in their risk of bias assessment conducted for systematic reviews.[17]

We also strongly recommend EPA identify tools that may potentially not be appropriate for human health chemical assessments. Many tools are currently being developed for evaluating risk of bias, quality, and strength of evidence for individual studies as well as for systematic reviews. It is critical that EPA evaluate tools developed in other fields that may be relevant, such as for clinical or preclinical animal or human studies, as these tools could potentially be modified for an application to questions of environmental health relevance. However, these tools should be applied with caution—due to the

---

[12] http://www.prisma-statement.org/; Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews. 2015 Jan 1;4(1):1.

[13] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

[14] NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[15] Bero LA. Why the Cochrane risk of bias tool should include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000075.*

Sterne JA. Why the Cochrane risk of bias tool should not include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000076.*

[16] Krauth D, Anglemyer A, Philipps R, Bero L. Nonindustry-sponsored preclinical studies on statins yield greater efficacy estimates than industry-sponsored studies: a meta-analysis. *PLoS Biol. 2014 Jan; 12(1):e1001770*

Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev. 2012 Dec 12; 12():MR000033.*

[17] NRC (National Research Council). Review of EPA's Integrated Risk Information System (IRIS) Process. Washington, DC:National Academies Press. 2014.

ED_004886_00001345-00009

differences in the types of evidence under evaluation a direct application to a difference evidence base than intended could lead to biased and erroneous conclusions.[18]

2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**

In general, EPA is proposing to consider three populations for exposure assessment: 1) Occupational users and non-users; 2) consumers and bystanders; and 3) general population. We strongly recommend that EPA calculate the aggregate exposures within and across these populations-- risk will be underestimated if it does not include these real-world exposures. Exposures within a population should also be aggregated (rather than considered in isolation) in order to estimate the general population's actual exposure to the chemical—for example, through exposures from food, water and air.

Further, as shown in the Figure below, exposures must also be aggregated across populations. Consumers and workers are part of the general population – that is, since workers and consumers also eat food and drink water, they will have the same exposures as the general population, in addition to the anticipated exposures on-the-job or from consumer products. Some workers will also be consumer product users, so they have the potential to face general, consumer product, and on-the-job exposures. These specific exposure scenarios must be accounted for in EPA's exposure estimation to ensure that such individual exposures are adequately considered and integrated into the risk assessment.

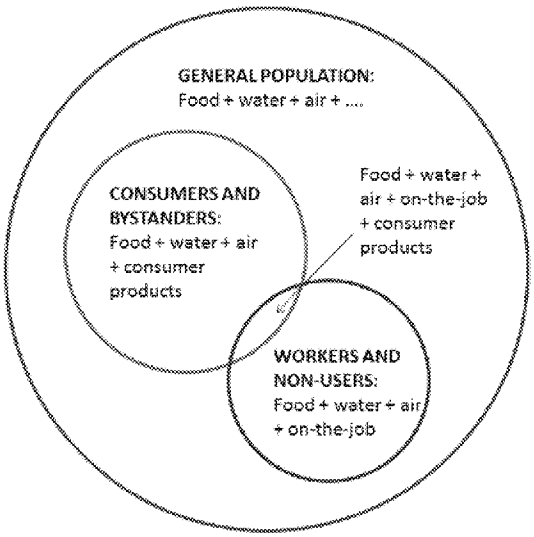


Figure: EPA must assess aggregate exposures within and across all the populations for accurate exposure assessment.

In the Introduction section of the chemical Scope documents, EPA states that it "may consider background exposures from legacy use, associated disposal, and legacy disposal as part of an assessment of aggregate exposure or as a tool to evaluate the risk of exposures resulting from non-

[18] Nachman, K.E., Lam, J., Schinasi, L.H., Smith, T.C., Feingold, B.J. and Casey, J.A., 2017. O'Connor et al. systematic review regarding animal feeding operations and public health: critical flaws may compromise conclusions. *Systematic Reviews*, *6*(1), p.179.

legacy uses."[19] This falls short of the analysis required under Lautenberg TSCA. It is critical that EPA consider ongoing exposures from legacy uses and disposal, and includes these as part of the aggregate exposure assessment. Asbestos and HBCD are two examples of this, as they have enormous volumes in place in buildings and existing infrastructure. The Healthy Building Network estimates there are 66 million- 132 million pounds (30,000-60,000 metric tons) of HBCD in insulation in existing buildings[20] — these reservoirs in-place are and will continue to be critical sources of ongoing exposures. HBCD was also used in cars and furniture, which are long-lived consumer items that will continue to contribute to ongoing exposures for years to come.

Another example is 1,4-dioxane, which was historically used as a chemical stabilizer for chlorinated solvents. Many groundwater aquifers are contaminated with 1,4-dioxane, and the extent of legacy contamination of groundwater is likely underestimated.[21] Also, 1,4-dioxane occurs in a wide variety of products including personal care products, detergents, waxes, and antifreeze, and 1,4-dioxane is a by-product in manufacturing processes involving ethylene oxide, such as the production of polyethylene terephthalate (PET), polyester, and surfactants. The use and disposal of 1,4-dioxane has led to past environmental contamination which contributes to on-going exposures.[22] The physical and chemical properties of 1,4-dioxane render it a persistent and highly mobile water contaminant: it is highly miscible in water.[23] Exposures via drinking water are documented back to the 1980s and continue today.[24] Results from EPA's Third Unregulated Contaminant Monitoring Rule (UCMR3) highlight that over 13% of 4,905 public drinking water systems serving >10,000 people had concentrations of 1,4-dioxane above the EPA Reference Concentration of 0.35 ppb 1,4-dioxane.[25] Furthermore, the UCMR3 results do not capture exposures in communities served by small public drinking water systems serving <10,000 people.[26] Approximately 27% of the US population is served by small public drinking water systems.[27] Thus, it will be critical for EPA to consider the population's current exposure to 1,4-dioxane via sources like drinking water as part of their assessment for health risks.

---

[19] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 12

[20] Safer Chemicals, Healthy Families et al. Comments to the U.S. Environmental Protection Agency (EPA) on the Scope of its Risk Evaluation for the TSCA Work Plan Chemicals: CYCLIC ALIPHATIC BROMIDE CLUSTER or HEXABROMOCYCLODODECANE (HBCD). March 15, 2017. https://healthybuilding.net/uploads/files/saferchemicals-hbcd.pdf

[21] David T. Adamson, Shaily Mahendra, Kenneth L. Walker, Jr., Sharon R. Rauch, Shayak Sengupta, and Charles J. Newell. A Multisite Survey To Identify the Scale of the 1,4-Dioxane Problem at Contaminated Groundwater Sites. *Environmental Science & Technology Letters* 2014 1 (5), 254-258 DOI: 10.1021/ez500092u

[22] Agency for Toxic Substances and Disease Registry. TOXICOLOGICAL PROFILE FOR 1,4-DIOXANE. Atlanta, GA; 2012.

[23] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

[24] Dietrich, A.M., D.S. Millington, and R.F. Christman. 1983. Specific identification of organic pollutants in Haw River water using gas chromatography/mass spectrometry. WRRI Report No. 206. North Carolina State University, Raleigh, NC.

[25] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

U.S. Environmental Protection Agency. Unregulated Contaminant Monitoring Rule 74 FR 51850, EPA-HQ-OW-2007-1189 FRL-8963-6. Washington, D.C.; 2009. https://federalregister.gov/a/E9-24287.s

[26] Knappe, D.R.U., Lopez-Velandia, C., Hopkins, Z., & Sun, M. (2016). Occurrence of 1,4-Dioxane in the Cape Fear River Watershed and Effectiveness of Water Treatment Options for 1,4-Dioxane Control. NC Water Resources Research Institute.

[27] EPA. 2017. Small Drinking Water Systems Research https://www.epa.gov/water-research/small-drinking-water-systems-research-0 (accessed September 17, 2017).

ED_004886_00001345-00011

When a chemical is present in products or media as a contaminant/ by-product, EPA needs to include and assess these exposures. We strongly recommend against ignoring or discounting these potential exposures routes. For example, EPA proposes to exclude from consideration conditions of use of 1,4-dioxane when it is present as contaminant in a wide variety of items, including household detergents, cosmetics/ toiletries, and foods.[28] This exclusion is not scientifically justified. Cosmetics and personal care products have the potential to contribute significantly to exposures, since people are applying them directly to their bodies, often multiple times per day, every day.

Finally, in the exposure assessments for methylene chloride, N-methylpyrrolidone and trichloroethylene, EPA is proposing to exclude uses it already assessed.[29] We agree that EPA does not need to re-assess these uses; these evaluations have been completed and finalized. However, unless and until such uses are banned, the exposures from these uses continue. Therefore, the new risk evaluations need to consider the contributions of these uses to exposures by using the exposure values from the previous assessments.

For the occupational exposure analysis plan, EPA states it will "Consider and incorporate applicable engineering controls and/or personal protective equipment into exposure scenarios."[30] However, these are not realistic assumptions nor are they appropriate for public health protection. EPA's own research shows that the primary factors influencing whether a user understands label information are the users' literacy and numeracy, which frequently correlate with the users' education and income.[31] Therefore, people with less education, lower income, and less advanced literary skills will be the most likely to not understand label instructions. These individuals already disproportionately bear the burden of exposures to multiple environmental hazards and the resulting health impacts; thereby placing further burden on this already stressed susceptible subpopulation. Further, appropriate personal protective equipment (PPE) for workers is often not provided by employers, or may not be fitted or working properly. When evaluating occupational exposures, EPA needs to take into consideration all potential and feasible routes of exposure, and should not exclude exposure routes based on assumptions of PPE and/ or exposure controls in place. These controls are not guaranteed and may change in the future, so to assume zero exposure via these routes would be inappropriate and a failure to adequately ensure health protections, especially for susceptible sub-populations as required by the Lautenberg TSCA.

In summary, EPA needs to account for all the sources of exposure or it will underestimate risk for all 10 chemicals. When analyzing aggregate exposures, "sentinel exposure" may be considered simultaneously, where appropriate. However, these are not mutually exclusive and EPA should not incorporate sentinel to the exclusion of aggregate.

3. EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations

---

[28] US EPA (2017). Scope of the Risk Evaluation for 1,4-dioxane. Pg. 21

[29] US EPA (2017). Scope of the Risk Evaluation for Methylene Chloride (Dichloromethane, DCM). Pg. 30
US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 19-20
US EPA (2017). Scope of the Risk Evaluation for Trichloroethylene. Pg. 27

[30] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 45

[31] US Environmental Protection Agency (EPA). 2016. "The Effectiveness of Labeling on Hazardous Chemicals and Other Products." Office of Chemical Safety and Pollution Prevention. RIN 2079-AK07.

ED_004886_00001345-00012

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

In general, EPA proposes to consider workers and occupational non-users, consumer and by-standers, and other groups within the general population in proximity to conditions of use as sub-populations who experience greater exposures. In particular, EPA has appropriately identified people who live or work near manufacturing, processing, distribution, use or disposal sites as facing greater exposures. Such communities are often low income and/ or people of color, exposed to a disproportionate share of pollution, environmental hazards, social and economic stressors. Multiple exposures to chemical and non-chemical stressors collectively increase the risk of harm, combined with synergistic effects with other health stressors in their daily lives such as limited access to quality health care.[32,33]

EPA's risk evaluation needs to fully account for the reality of cumulative exposures, as recommended by the NAS in their Phthalates and Cumulative Risk report.[34] As described below, EPA can use "default values" to account for cumulative exposures.

In regards to greater susceptibility, EPA's considerations for addressing susceptibility vary considerably across the 10 chemicals. EPA should apply a consistent approach to addressing susceptibility across the 10 chemicals. The following are well-known factors that increase biologic sensitivity or reduce resilience to exposures,[35,36] and these as well as other relevant factors should be standard considerations for all 10 chemicals to identify susceptible sub-populations:

Intrinsic/ endogenous factors

- Genetic polymorphisms/ genetics/ genetic makeup
- Health status/ nutritional status/ disease status/ pre-existing conditions
- Prenatal lifestage
- Age

Extrinsic factors

- Multiple exposures/ co-exposures
- Race/ ethnicity
- Socioeconomic status (SES)

For example, the prenatal lifestage is the most sensitive to developmental and reproductive toxicants, and women of child-bearing age should be considered as a susceptible sub-population for any chemicals with such hazards. Women of reproductive age are not specifically identified as a potential susceptible

---

[32] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[33] Vesterinen HM, Morello-Frosch R, Sen S, Zeise L, Woodruff TJ. Cumulative effects of prenatal-exposure to exogenous chemicals and psychosocial stress on fetal growth: Systematic-review of the human and animal evidence. Meliker J, editor. PLoS One. 2017 Jul 12;12(7):e0176331.

[34] National Research Council. Committee on the Health Risks of Phthalates, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies. 2008. Phthalates and cumulative risk assessment: the task ahead. Washington, D.C.: National Academies Press.

[35] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[36] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

ED_004886_00001345-00013

sub-population for pigment violet 29, TCE, NMP, PERC, or HBCD, even though EPA will consider reproductive and developmental toxicity hazards for these chemicals.

As discussed below, science-based defaults should be used to account for these and other susceptibilities, unless there is there is chemical-specific data available to support increasing or decreasing the default.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**

EPA cites existing IRIS assessments for five chemicals; because these are EPA's own assessments which have gone through the Agency's peer-review process, and in some cases NAS review, EPA can rely on these existing finalized, authoritative assessments for hazard identification.

Moving forward, a weight of evidence evaluation is required by law, which EPA defines as:

> "Weight of scientific evidence means a systematic review method, applied in a manner suited to the nature of the evidence or decision, that uses a pre-established protocol to comprehensively, objectively, transparently, and consistently, identify and evaluate each stream of evidence, including strengths, limitations, and relevance of each study and to integrate evidence as necessary and appropriate based upon strengths, limitations, and relevance."[37]

Therefore, EPA should use a systematic review process for evaluating scientific information for chemicals that do not have an IRIS assessment and for any additional studies that will be considered for the chemicals that have IRIS assessments.

For the scoping document, EPA should include all hazards identified in the literature, and not make decisions about their relevance to the risk evaluation until a systematic review has been completed. For a number of chemicals, EPA has inappropriately drawn conclusions about hazards prior to the completion of a systematic review. Some examples are given in the table below where EPA concludes that HBCD, NMP and pigment violet 29 are not genotoxic based on previous assessments and without conducting a systematic review.

Chemical	Example Text from EPA Scoping Document
HBCD	"Available data suggest that HBCD is not genotoxic. Existing assessments have also concluded, based on genotoxicity information and a limited lifetime study, that HBCD is not carcinogenic (NICNAS, 2012; EINECS, 2008; TemaNord, 2008; OECD, 2007). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity or cancer hazards in the risk evaluation of HBCD at this time."[38]
NMP	"NMP is not mutagenic, based on results from bacterial and mammalian *in vitro* tests and *in vivo* systems and is not considered to be carcinogenic (RIVM, 2013; OECD,

[37] 82 Fed. Reg. 138, 33748
[38] US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg.36

	2007; WHO, 2001). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity and cancer hazards in the NMP risk evaluation."[39]
Pigment violet 29	"Testing for carcinogenicity of Pigment Violet 29 has not been conducted. However, negative genotoxicity results, structure-activity considerations and the expectation of negligible absorption and uptake of Pigment Violet 29 (based on very low solubility), indicate carcinogenicity of Pigment Violet 29 is unlikely. Unless new information indicates otherwise, EPA does not expect to conduct additional, in-depth analyses of genotoxicity and cancer hazards in the risk evaluation of Pigment Violet 29."[40]

The National Academies recently released a report with recommendations on implementation of systematic review for EPA's chemical evaluations (which we will refer to as the 'NAS Systematic Review report' for simplicity).[41] First, they recommend that EPA should develop policies and procedures that allow the agency to use and update existing systematic reviews, since the committee concluded that could potentially save time and resources. EPA should conduct a review to determine whether there are existing systematic reviews on the topic of interest and if there is, EPA should evaluate it to determine if it is high-quality. The NAS recommends that EPA should build on existing high-quality reviews to incorporate new studies and use the updated systematic review as a basis for its assessment. The assessments cited by EPA to support the hazard identification claims are not systematic reviews; even if they were, EPA should evaluate them for quality before relying on their conclusions.

Second, it is very likely that additional studies have been published since the assessments EPA cites were completed. EPA should develop criteria to evaluate the internal validity (risk of bias) of individual studies, utilizing existing tools that have been developed and empirically demonstrated on environmental health studies such as the Navigation Guide or the OHAT approach.[42] We also recommend that EPA not using a scoring system to evaluate study quality. Specifically, we note that empirically validated approaches in the clinical sciences such as Cochrane discourage using a numerical scale scoring approach for evaluating study quality because calculating a score requires choosing a weighting scheme for each component, which generally is nearly impossible to justify.[43] Furthermore, a study might be well designed to eliminate bias, but because the study failed to report details in the publication under review, it will receive a low score--most available scoring systems include a mix of risk of bias and reporting biases which is inappropriate. Additionally, quality scores have been shown to be invalid for assessing risk of bias in clinical research.[44] The current standard in evaluation of both clinical

[39] US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 36

[40] US EPA (2017). Scope of the Risk Evaluation for Pigment Violet 29. Pg.29

[41] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[42] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[43] Juni, P., A. Witschi, R. Bloch, and M. Egger. 1999. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 282(11):1054-1060.

[44] Id.

and environmental health research calls for reporting each component of the assessment tool separately and not calculating an overall numeric score.[45]

Data generated by alternative test methods (such as high-throughput screening methods) are not different than any other type of *in vitro* or cell-based assay data that would be considered in a systematic review. These kinds of assays provide mechanistic data, and the NAS Systematic Review report explicitly considered how mechanistic data could be utilized in a systematic review for evidence integration. The committee came to two conclusions. First, the same protocol for evaluating relevance and study quality must be used with mechanistic data as for any other study. For example, in the report's case study on phthalates, the committee was not able to integrate results from high-throughput assays because the cell lines used were of unknown relevance to the *in vivo* mechanism of phthalate toxicity.[46] Second, the foundation of the hazard classification in a systematic review is the animal and human data, with the mechanistic data playing a supporting role. If mechanistic data is relevant, it can be used to upgrade a hazard classification, or increase the confidence of a finding made based on evaluation of animal and human data. A hazard classification is never made based on high-throughput or other kinds of mechanistic data alone.[47]

5. For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.

Defaults

We strongly support the use of health protective defaults to incorporate factors that reflect the range of variability and susceptibility in the population to ensure risks are not underestimated. The importance of using protective science-based defaults was highlighted by the NAS in 2009.[48] The default should be used for factors that are known to influence risk unless there is chemical-specific data that support increasing or decreasing it; when there is inadequate information to quantitatively assess inter- or intra-species differences for a specific chemical, the defaults should be used. For example, EPA's defaults should include:

- Inter-human variability, general
- Inter-human susceptibility to carcinogens, adult
- Inter-human susceptibility to carcinogens, early life (including prenatal)
- Inter-human susceptibility to non-carcinogens, early life (including prenatal)
- Animal findings are relevant to humans
- Findings from one route of exposure are considered representative unless data show otherwise

---

[45] Higgins, J.P.T., and S. Green, eds. 2008. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

[46] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017. Pg. 78

[47] Id. Pp. 158-9

[48] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch 4-6

ED_004886_00001345-00016

EPA has relied on standard default values ("uncertainty" or "safety" factors) that have been applied across the board to various chemicals and health outcomes. But newer science demonstrates that EPA's typical safety factor of 10 is insufficient to account for variability due to life stage, genetics, underlying disease status, and external stressors that may be due to poverty or other difficult life conditions.

For cancer, the NAS recommended that EPA include a factor to account for human variability in response to carcinogens, as EPA's current approach inaccurately assumes that there is no variability in response. They found that a factor of 25- to 50- may account for the variability between the median individual and those with more extreme responses, and recommended 25 as a reasonable default value.[49]

Similarly, EPA should increase or add factors that address cancer and non-cancer susceptibility during early life stages. While EPA does account for increased susceptibility to genotoxicants, it does not include the prenatal period or chemicals that can influence cancer through other mechanisms. California EPA's guidance incorporates factors to account for increased susceptibility for exposures that occur prenatally for carcinogens, non-mutagenic carcinogenic agents and non-carcinogens. Their literature review on differential susceptibility to carcinogens and non-carcinogens based on age and life stage derived age adjustment values for carcinogens which include the prenatal period[50] and increased the default intraspecies uncertainty factors for non-carcinogens to 30 and 100 for specific endpoints such as asthma or neurotoxicity.[51] At a minimum, EPA should use Cal EPA's age adjustment values and intraspecies uncertainty factors for incorporating age/early life susceptibility.

In general, developmental life stages, including the fetus, infancy, and childhood, are more vulnerable to chemical exposure and toxicity. However, typical EPA age-dependent adjustment factors account for other life stages but NOT fetal exposures. Recent studies have demonstrated differential expression and activity of metabolic enzymes such as Cytochrome P450 in fetal versus adult tissue, indicating potential lifestage-dependent variability in metabolic capabilities and greater vulnerability during fetal development not accounted for in current risk assessment practices.[52] This is a critical point to address, as disruptions during fetal development have implications for health and disease in adulthood. EPA should evaluate this rich body of literature to identify the most up-to-date scientific knowledge regarding human variability and susceptibility and incorporate these scientifically-based default values in their assessments unless there are chemical-specific data supporting departing from the defaults. California EPA also developed child-specific risk values for chemicals (e.g., atrazine, lead, nickel, manganese, heptachlor) that specifically address routes of exposure and differences in susceptibility

---

[49] Id. Pg. 168

[50] California EPA 2009. Cal EPA 2009. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document for Cancer Potency Factors: Methodologies for derivation, listing of available values, and adjustments to allow for early life stage exposures. http://oehha.ca.gov/media/downloads/crnr/tsdcancerpotency.pdf

[51] Cal EPA 2008. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document For the Derivation of Noncancer Reference Exposure Levels http://oehha.ca.gov/media/downloads/crnr/noncancertsdfinal.pdf

[52] Sadler, N.C., Nandhikonda, P., Webb-Robertson, B.J., Ansong, C., Anderson, L.N., Smith, J.N., Corley, R.A. and Wright, A.T., 2016. Hepatic cytochrome P450 activity, abundance, and expression throughout human development. *Drug Metabolism and Disposition*, *44*(7), pp.984-991.

ED_004886_00001345-00017

unique to children compared to adults.[53] EPA should review these evaluations and incorporate these values as appropriate. Furthermore, a default guidance principle should be that animal findings are relevant to humans unless there is sufficient and compelling information to support otherwise.

Risk estimates

EPA should not use MOE (margin of exposure) as an analysis method in the risk evaluation process moving forward. MOE is not an estimate of risk—it is a single number that is a version of the "bright line" approach like the Reference Dose (or Reference Concentration for inhalation doses). MOE is calculated by dividing the point of departure (e.g., LOAELs, NOAELs or BMDLs) by estimated exposure values, and this 'bright line' approach does not provide information about the magnitude of the risks above, at, or below this line. Further, it implies that there is a "safe" level of exposure below which no harm will occur. While this may be true for a select few chemicals, the NAS Science and Decisions report recognizes that this is not a valid assumption for all chemicals and has recommended moving away from such "bright line" approaches which do not establish risk estimates across the full range of exposures.[54] Additionally, the MOE will not provide the necessary information for future analysis of risks and benefits that will be critical for decision-making on these chemicals.[55] We recommend that EPA utilize available analytical methods such as PODs based on a BMD to develop quantified estimates of risk.

EPA appropriately states that a dose-response assessment will be conducted for all identified human health hazard endpoints. PODs should also be developed for every endpoint unless the data are insufficient to develop a model. For calculating cancer or non-cancer risks, we recommend always using a point of departure (POD) of a benchmark dose (BMD) at 1%. The POD should be based on a BMD calculation, not the NOAEL/LOAEL, unless the data are insufficient to model. EPA already recognizes the features that make BMDs superior: BMDs account for the shape of the dose–response function; are independent of study design, such as the space between dosing; and are comparable across chemicals.[56]

Historically, for carcinogens that are direct mutagens or are associated with large human body burdens, EPA has assumed there is no threshold of effect. But the NAS Science and Decisions report highlights the science indicating that this linear presumption with no threshold is appropriate for the calculation of both cancer and non-cancer risks, and regardless of whether a carcinogen is a mutagen. For example, dose-response relationships can be linear at low dose when exposures contribute to an existing disease process, add to background processes and/ or exposures, and interact with interindividual variability or susceptibility.[57] Science and Decisions recommends harmonizing cancer and non-cancer risk assessment approaches. Therefore, for calculating non-mutagen cancer or non-cancer risks based on a POD, EPA should use the same approach as for mutagens, which assumes a straight line from the POD. In fact, a

---

[53] California Environmental Protection Agency. Office of Environmental Health Hazard Assessment (OEHHA). Child-Specific Reference Doses (chRDs) Finalized to Date. Available from: http://oehha.ca.gov/risk-assessment/chrd/table-all-chrds

[54] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

[55] McGartland A, Revesz R, Axelrad DA, Dockins C, Sutton P, Woodruff TJ. Estimating the health benefits of environmental regulations. Science (80-). 2017 Aug 4;357(6350):457–8.

[56] Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, Rusyn I. 2014. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. Environmental health perspectives. 122(5).

[57] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch.5

ED_004886_00001345-00018

linear relationship may actually underestimate risks for some chemicals where the dose-response curve is supra-linear.

6. Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.

Production volumes for both asbestos and pigment violet 29 have been claimed as CBI. Production volume is basic information about a chemical to which the public and scientists should have access. We urge EPA to move forward with substantiating such claims under the new TSCA.

Appendix

When and how to update systematic reviews: consensus and checklist
Garner et al.
BMJ 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

ED_004886_00001345-00020

OPEN ACCESS

When and how to update systematic reviews: consensus and checklist

Paul Garner,[1] Sally Hopewell,[2] Jackie Chandler,[3] Harriet MacLehose,[3] Elie A Akl,[5,6] Joseph Beyene,[7] Stephanie Chang,[8] Rachel Churchill,[9] Karin Dearness,[10] Gordon Guyatt,[4] Carol Lefebvre,[11] Beth Liles,[12] Rachel Marshall,[3] Laura Martínez García,[13] Chris Mavergames,[14] Mona Nasser,[15] Amir Qaseem,[16,17] Margaret Sampson,[18] Karla Soares-Weiser,[3] Yemisi Takwoingi,[19] Lehana Thabane,[4,20] Marialena Trivella,[21] Peter Tugwell,[22] Emma Welsh,[23] Ed C Wilson,[2,4] Holger J Schünemann[4,5]

For numbered affiliations see end of article.

Correspondence to: P Garner Paul.Garner@lstmed.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

Accepted: 26 May 2016

Updating of systematic reviews is generally more efficient than starting all over again when new evidence emerges, but to date there has been no clear guidance on how to do this. This guidance helps authors of systematic reviews, commissioners, and editors decide when to update a systematic review, and then how to go about updating the review.

Systematic reviews synthesise relevant research around a particular question. Preparing a systematic review is time and resource consuming, and provides a snapshot of knowledge at the time of incorporation of data from studies identified during the latest search. Newly identified studies can change the conclusion of a review. If they have not been included, this threatens the validity of the review, and, at worst, means the review could mislead. For patients and other healthcare consumers, this means that care and policy development might not be fully informed by the latest research; furthermore, researchers could be misled and carry out research in areas where no further research is actually needed.[1] Thus, there are clear benefits to updating reviews, rather than duplicating the entire process as new evidence emerges or new methods develop. Indeed, there is probably added value to updating a review, because this will include taking into account comments and criticisms, and adoption of new methods in an iterative process.[2-6]

Cochrane has over 20 years of experience with preparing and updating systematic reviews, with the publication of over 6000 systematic reviews. However, Cochrane's principle of keeping all reviews up to date has not been possible, and the organisation has had to adapt: from updating when new evidence becomes available,[7] to updating every two years,[8] to updating based on need and priority.[9] This experience has shown that it is not possible, sensible, or feasible

SUMMARY POINTS

Updating systematic reviews is, in general, more efficient than starting afresh when new evidence emerges. The panel for updating guidance for systematic reviews (PUGs; comprising review authors, editors, statisticians, information specialists, related methodologists, and guideline developers) met to develop guidance for people considering updating systematic reviews. The panel proposed the following:

1. Decisions about whether and when to update a systematic review are judgments made for individual reviews at a particular time. These decisions can be made by agencies responsible for systematic review portfolios, journal editors with systematic review update services, or author teams considering embarking on an update of a review.
2. The decision needs to take into account whether the review addresses a current question, uses valid methods, and is well conducted; and whether there are new relevant methods, new studies, or new information on existing included studies. Given this information, the agency, editors, or authors need to judge whether the update will influence the review findings or credibility sufficiently to justify the effort in updating it.
3. Review authors and commissioners can use a decision framework and checklist to navigate and report these decisions with "update status" and rationale for this status. The panel noted that the incorporation of new synthesis methods (such as Grading of Recommendations Assessment, Development and Evaluation (GRADE)) is also often likely to improve the quality of the analysis and the clarity of the findings.
4. Given a decision to update, the process needs to start with an appraisal and revision of the background, question, inclusion criteria, and methods of the existing review.
5. Search strategies should be refined, taking into account changes in the question or inclusion criteria. An analysis of yield from the previous edition, in relation to databases searched, terms, and languages can make searches more specific and efficient.
6. In many instances, an update represents a new edition of the review, and authorship of the new version needs to follow criteria of the International Committee of Medical Journal Editors (ICMJE). New approaches to publishing licences could help new authors build on and re-use the previous edition while giving appropriate credit to the previous authors.

The panel also reflected on this guidance in the context of emerging technological advances in software, information retrieval, and electronic linkage and mining. With good synthesis and technology partnerships, these advances could revolutionise the efficiency of updating in the coming years.

ED_004886_00001345-00021

to continually update all reviews all the time. Other groups, including guideline developers and journal editors, adopt updating principles (as applied, for example, by the *Systematic Reviews* journal; https://systematicreviewsjournal.biomedcentral.com/).

The panel for updating guidance for systematic reviews (PUGs) group met to draw together experiences and identify a common approach. The PUGs guidance can help individuals or academic teams working outside of a commissioning agency or Cochrane, who are considering writing a systematic review for a journal or to prepare for a research project. The guidance could also help these groups decide whether their effort is worthwhile.

Panel selection and procedures

An international panel of authors, editors, clinicians, statisticians, information specialists, other methodologists, and guideline developers was invited to a two day workshop at McMaster University, Hamilton, Canada, on 26-27 June 2014, organised by Cochrane. The organising committee selected the panel (web appendix 1). The organising committee invited participants, put forward the agenda, collected background materials and literature, and drafted the structure of the report.

The purpose of the workshop was to develop a common approach to updating systematic reviews, drawing on existing strategies, research, and experience of people working in this area. The selection of participants aimed on broad representation of different groups involved in producing systematic reviews (including authors, editors, statisticians, information specialists, and other methodologists), and those using the reviews (guideline developers and clinicians). Participants within these groups were selected on their expertise and experience in updating, in previous work developing methods to assess reviews, and because some were recognised for developing approaches within organisations to manage updating strategically. We sought to identify general approaches in this area, and not be specific to Cochrane; although inevitably most of the panel were somehow engaged in Cochrane.

The workshop structure followed a series of short presentations addressing key questions on whether, when, and how to update systematic reviews. The proceedings included the management of authorship and editorial decisions, and innovative and technological approaches. A series of small group discussions followed each question, deliberating content, and forming recommendations, as well as recognising uncertainties. Large group, round table discussions deliberated further these small group developments. Recommendations were presented to an invited forum of individuals with varying levels of expertise in systematic reviews from McMaster University (of over 40 people), widely known for its contributions to the field of research evidence synthesis. Their comments helped inform the emerging guidance.

The organising committee became the writing committee after the meeting. They developed the guidance arising from the meeting, developed the checklist and diagrams, added examples, and finalised the manuscript. The guidance was circulated to the larger group three times, with the PUGs panel providing extensive feedback. This feedback was all considered and carefully addressed by the writing committee. The writing committee provided the panel with the option of expressing any additional comments from the general or specific guidance in the report, and the option for registering their own view that might differ to the guidance formed and their view would be recorded in an annex. In the event, consensus was reached, and the annex was not required.

Definition of update

The PUGs panel defined an update of a systematic review as a new edition of a published systematic review with changes that can include new data, new methods, or new analyses to the previous edition. This expands on a previous definition of a systematic review update.[10] An update asks a similar question with regard to the participants, intervention, comparisons, and outcomes (PICO) and has similar objectives; thus it has similar inclusion criteria. These inclusion criteria can be modified in the light of developments within the topic area with new interventions, new standards, and new approaches. Updates will include a new search for potentially relevant studies and incorporate any eligible studies or data; and adjust the findings and conclusions as appropriate. Box 1 provides some examples.

Which systematic reviews should be updated and when?

Any group maintaining a portfolio of systematic reviews as part of their normative work, such as guidelines panels or Cochrane review groups, will need to prioritise which reviews to update. Box 2 presents the approaches used by the Agency for HealthCare Research and Quality (AHRQ) and Cochrane to prioritise which systematic reviews to update and when. Clearly, the responsibility for deciding which systematic reviews should be updated and when they will be updated will vary: it may be centrally organised and resourced, as with the AHRQ scientific resource centre (box 2). In Cochrane, the decision making process is decentralised to the Cochrane Review Group editorial team, with different approaches applied, often informally.

Box 1: Examples of what factors might change in an updated systematic review

- A systematic review of steroid treatment in tuberculosis meningitis used GRADE methods and split the composite outcome in the original review of death plus disability into its two components. This improved the clarity of the reviews findings in relation to the effects and the importance of the effects of steroids on death and on disability.[11]
- A systematic review of dihydroartemisinin-piperaquine (DHAP) for treating malaria was updated with much more detailed analysis of the adverse effect data from the existing trials as a result of questions raised by the European Medicines Agency. Because the original review included other comparisons, the update required extracting only the DHAP comparisons from the original review, and a modification of the title and the PICO.[12]
- A systematic review of atorvastatin was updated with simple uncontrolled studies.[13] This update allowed comparisons with trials and strengthened the review findings.[14]

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001345-00022

The PUGs panel recommended an individualised approach to updating, which used the procedures summarised in figure 1. The figure provides a status category, and some options for classifying reviews into each of these categories, and builds on a previous decision tool and earlier work developing an updating classification system.[15 16] We provide a narrative for each step.

Step 1: assess currency

Does the published review still address a current question?

An update is only worthwhile if the question is topical for decision making for practice, policy, or research priorities (fig 1). For agencies, people responsible for managing a portfolio of systematic reviews, there is a need to use both formal and informal horizon scanning. This type of scanning helps identify questions with currency, and can help identify those reviews that should be updated. The process could include monitoring policy debates around the review, media outlets, scientific (and professional) publications, and linking with guideline developers.

Has the review had good access or use?

Metrics for citations, article access and downloads, and sharing via social or traditional media can be used as proxy or indicators for currency and relevance of the review. Reviews that are widely cited and used could be important to update should the need arise. Comparable reviews that are never cited or rarely downloaded, for example, could indicate that they are not addressing a question that is valued, and might not be worth updating.

In most cases, updated reviews are most useful to stakeholders when there is new information or methods that result in a change in findings. However, there are some circumstances in which an up to date search for information is important for retaining the credibility of the review, regardless of whether the main findings would change or not. For example, key stakeholders would dismiss a review if a study is carried out in a relevant geographical setting but is not included; if a large, high profile study that might not change the findings is not included; or if an up to date search is required for a guideline to achieve credibility. Box 3 provides such examples. If the review does not answer a current question, the intervention has been superseded, then a decision can be made not to update and no further intelligence gathering is required (fig 1).

Did the review use valid methods and was it well conducted?

If the question is current and clearly defined, the systematic review needs to have used valid methods and be well conducted. If the review has vague inclusion criteria, poorly articulated outcomes, or inappropriate methods, then updating should not proceed. If the question is current, and the review has been cited or used, then it might be appropriate to simply start with a new protocol. The appraisal should take into account the methods in use when the review was done.

Step 2: identify relevant new methods, studies, and other information

Are there any new relevant methods?

If the question is current, but the review was done some years ago, the quality of the review might not meet current day standards. Methods have advanced quickly, and data extraction and understanding of the review process have become more sophisticated. For example:

- Methods for assessing risk of bias of randomised trials,[23] diagnostic test accuracy (QUADAS-2),[24] and observational studies (ROBINS-1).[25]
- Application of summary of findings, evidence profiles, and related GRADE methods has meant the characteristics of the intervention, characteristics of the participants, and risk of bias are more thoroughly and systematically documented.[26 27]
- Integration of other study designs containing evidence, such economic evaluation and qualitative research.[28]

There are other incremental improvements in a wide range of statistical and methodological areas, for example, in describing and taking into account cluster

Box 2: Examples of how different organisations decide on updating systematic reviews

Agency for Healthcare Research and Quality (US)

The AHRQ uses a needs based approach; updating systematic reviews depends on an assessment of several criteria:

1. Stakeholder impact
   - Interest from stakeholder partners (such as consumers, funders, guideline developers, clinical societies, James Lind Alliance)
   - Use and uptake (for example, frequency of citations and downloads)
   - Citation in scientific literature including clinical practice guidelines
2. Currency and need for update
   - New research is available
   - Review conclusions are probably dated
3. Update decision
   - Based on the above criteria, the decision is made to either update, archive, or continue surveillance.

Cochrane

Of over 50 Cochrane editorial teams, most but not all have some systems for updating, although this process can be informal and loosely applied. Most editorial teams draw on some or all of the following criteria:

1. Strategic importance
   - Is the topic a priority area (for example, in current debates or considered by guidelines groups)?
   - Is there important new information available?
2. Practicalities in organising the update that many groups take into account
   - Size of the task (size and quality of the review, and how many new studies or analyses are needed)
   - Availability and willingness of the author team
3. Impact of update
   - New research impact on findings and credibility
   - Consider whether new methods will improve review quality
4. Update decision
   - Priority to update, postpone update, class review as no longer requiring an update

ED_004886_00001345-00023

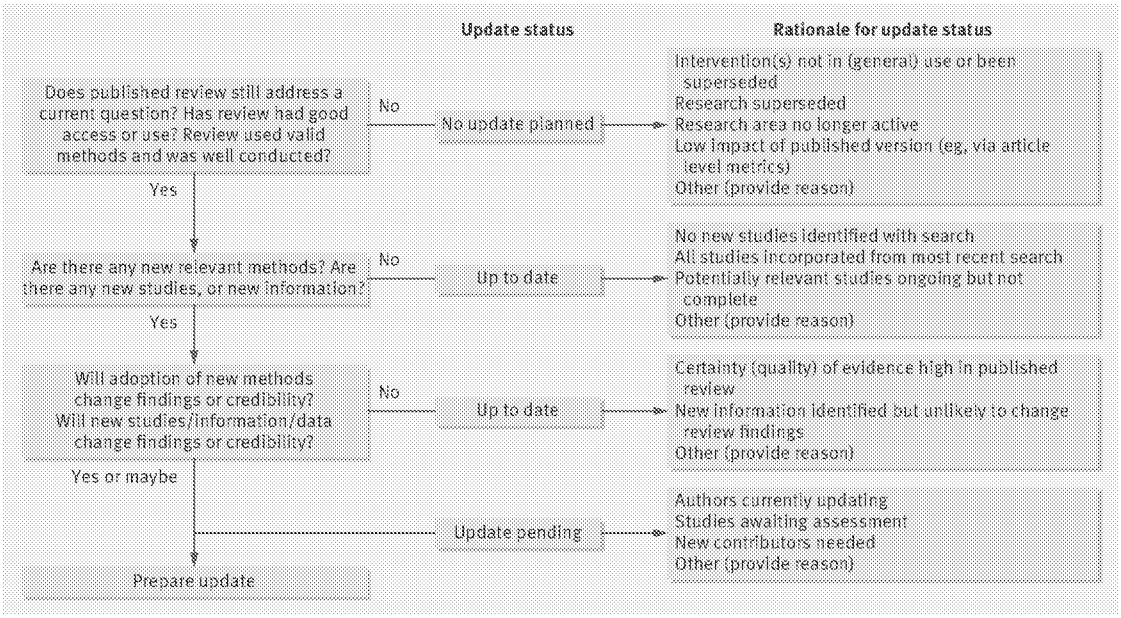


Fig 1 | Decision framework to assess systematic reviews for updating, with standard terms to report such decisions

randomised trials.[29] AMSTAR can assess the overall quality of a systematic review,[30] and the ROBIS tool can provide a more detailed assessment of the potential for bias.[31]

Are there any new studies or other information?

If an authoring or commissioning team wants to ensure that a particular review is up to date, there is a need for routine surveillance for new studies that are potentially relevant to the review, by searching and trial register inspection at regular intervals. This process has several approaches, including:

- Formal surveillance searching[32]
- Updating the full search strategies in the original review and running the searches
- Tracking studies in clinical trial and other registers
- Using literature appraisal services[33]
- Using a defined abbreviated search strategy for the update[34]
- Checking studies included in related systematic reviews.[35]

How often this surveillance is done, and which approaches to use, depend on the circumstances and the topic. Some topics move quickly, and the definition of "regular intervals" will vary according to the field and according to the state of evidence in the field. For example, early in the life of a new intervention, there might be a plethora of studies, and surveillance would be needed more frequently.

Box 3: Examples of a systematic review's currency

- The public is interested in vitamin C for preventing the common cold: the Cochrane review includes over 29 trials with either no or small effects, concluding good evidence of no important effects.[17] Assessment: still a current question for the public.
- Low osmolarity oral rehydration salt (ORS) solution versus standard solution for acute diarrhoea in children: the 2001 Cochrane review[18] led the World Health Organization to recommend ORS solution formula worldwide to follow the new ORS solution formula[19] and this has now been accepted globally. Assessment: no longer a current question.
- Routine prophylactic antibiotics with caesarean section: the Cochrane review reports clear evidence of maternal benefit from placebo controlled trials but no information on the effects on the baby.[20] Assessment: this is a current question.
- A systematic review published in the *Lancet* examined the effects of artemisinin based combination treatments compared with monotherapy for treating malaria and showed clear benefit.[21] Assessment: this established the treatment globally and is no longer a current question and no update is required.
- A Cochrane review of amalgam restorations for dental caries[22] is unlikely to be updated because the use of dental amalgam is declining, and the question is not seen as being important by many dental specialists. Assessment: no longer a current question.

Step 3: assess the effect of updating the review

Will the adoption of new methods change the findings or credibility?

Editors, referees, or experts in the topic area or methodologists can provide an informed view of whether a review can be substantially improved by application of current methodological expectations and new methods (fig 1). For example, a Cochrane review of iron supplementation in malaria concluded that there was "no significant difference between iron and placebo detected."[36] An update of the review included a GRADE assessment of the certainty of the evidence, and was able to conclude with a high degree of certainty that iron does not cause an excess of clinical malaria because the upper relative risk confidence intervals of harm was 1.0 with high certainty of evidence.[37]

Will the new studies, information, or data change the findings or credibility?

The assessment of new data contained in new studies and how these data might change the review is often used to determine whether an update should go ahead,

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001345-00024

Table 1 | Formal prediction tools: how potentially relevant new studies can affect review conclusions

Method	Description of approach	How it could be used	Advantages	Limitations	Validation
GRADE approach[38]	Considers whether the evidence certainty might change in the update (for example, because of lack of high certainty evidence, or because new evidence contradicts existing high certainty evidence). High certainty of evidence for critical outcomes could lower the priority for updating. Uncertainty in the review findings increases the need to include new studies[38]	Provides a benchmark by outcome to assess whether a new trial will improve the certainty of the evidence	Pragmatic. Many reviews already include GRADE	Requires GRADE to have been used in existing review or to complete an assessment according to GRADE	GRADE summary of findings tables or evidence profiles widely validated. Use of GRADE approach to prioritising updates requires further validation
Ottawa method[39-41]	A simple PubMed search (using the three largest and three most recent trials from the original review) to identify new studies. If new studies are found, then the method uses quantitative signals (eg, change in significance, effect size) to assess the likelihood that the new studies will change the review conclusion, thus triggering an update	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions; reasonable agreement with RAND method[39 42-44]
RAND method[45]	An abbreviated search of five major journals to identify new studies, and a search of the US Food and Drug Administration website and external expert judgment to determine the currency of the report findings	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions, and compares well with the Ottawa method[39 43 44]
Statistical prediction tool[15]	A multicomponent decision tool to assess whether there might be any new studies for the update. If new studies are identified, a statistical prediction tool estimates the probability that this will change the review conclusion	Ranks multiple systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires commercial software	Approach validated internally[46]; requires further external validation
Value of information analysis[47 48]	Builds on the statistical prediction tool approach[15] comparing the expected health gain from new evidence with its expected cost. The gain is calculated in terms of a reduction in expected loss from reduced uncertainty and the cost is measured in days required to update the review.	Ranks selected systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires some statistical knowledge	Approach validated internally; requires further external validation

Box 4: Examples of new information other than new trials being important

- The iconic Cochrane review of steroids in preterm labour was thought to provide evidence of benefit in infants, and this question no longer required new trials. However, a new large trial published in the *Lancet* in 2015 showed that in low and middle income countries, strategies to promote the uptake of neonatal steroids increased neonatal mortality and suspected maternal infection.[49] This information needs to somehow be incorporated into the review to maintain its credibility.
- A Cochrane review of community deworming in developing countries indicates that in recent studies, there is little or no effect.[50] The inclusion of a large trial of two million children confirmed that there was no effect on mortality. Although the incorporation of the trial in the review did not change the review's conclusions, the trial's absence would have affected the credibility of the review, so it was therefore updated.
- A new paper reporting long term follow-up data on anthracycline chemotherapy as part of cancer treatment was published. Although the effects from the outcomes remained essentially unchanged, apart from this longer follow-up, the paper also included information about the performance bias in the original trial, shifting the risk of bias for several outcomes from "unknown" to "high" in the Cochrane review.[51]

and the speed with which the update should be conducted. The appraisal of these new data can be carried out in different ways. Initially, methods focused on statistical approaches to predict an overturning of the current review findings in terms of the primary or desired outcome (table 1). Although this aspect is important, additional studies can add important information to a review, which is more than just changing the primary outcome to a more accurate and reliable estimate. Box 4 gives examples.

Reviews with a high level of certainty in the results (that is, when the GRADE assessment for the body of evidence is high) are less likely to change even with the addition of new studies, information, or data, by definition. GRADE can help guide priorities in whether to update, but it is still important to assess new studies that might meet the inclusion criteria. New studies can show unexpected effects (eg, attenuation of efficacy) or provide new information about the effects seen in different circumstances (eg, groups of patients or locations).

Other tools are specifically designed to help decision making in updating. For example, the Ottawa[39] and RAND[45] methods focus on identification of new evidence, the statistical predication tool[15] calculates the probability of new evidence changing the review conclusion, and the value of information analysis approach[52] calculates the expected health gain (table 1). As yet, there has been limited external validation of these tools to determine which approach would be most effective and when.

If potentially relevant studies are identified that have not previously been assessed for inclusion, authors or those managing the updating process need to assess whether including them might affect the conclusions of the review. They need to examine the weight and certainty of the new evidence to help determine whether an update is needed and how urgent that update is. The updating team can assess this informally by judging whether new studies or data are likely to substantively affect the review, for example, by altering the certainty in an existing comparison, or by generating new comparisons and analyses in the existing review.

New information can also include fresh follow-up data on existing included studies, or information on how the studies were carried out. These should be assessed in terms of whether they might change the review findings or improve its credibility (fig 1). Indeed, if any study has been retracted, it is important the authors assess the reasons for its retraction. In the case of data fabrication, the study needs to be removed

ED_004886_00001345-00025

from the analysis and this recorded. A decision needs to be made as to whether other studies by the same author should be removed from the review and other related reviews. An investigation should also be initiated following guidelines from the Committee on Publication Ethics (COPE). Additional published and unpublished data can become available from a wide range of sources—including study investigators, regulatory agencies and industry—and are important to consider.

Preparing for an update

Refresh background, objectives, inclusion criteria, and methods

Before including new studies in the review, authors need to revisit the background, objectives, inclusion criteria, and methods of the current review. In Cochrane, this is referred to as the protocol, and editors are part of this process. The update could range from simply endorsing the current question and inclusion criteria, through to full rewriting of the question, inclusion criteria and methods, and republishing the protocol. As a field progresses with larger and better quality trials rigorously testing the questions posed, it may be appropriate to exclude weaker study designs (such as quasi-randomised comparisons or very small trials) from the update (table 2). The PUGs panel recommended that a protocol refresh will require the authors to use the latest accepted methods of synthesis, even if this means repeating data extraction for all studies.

New authors and authorship

Updated systematic reviews are new publications with new citations. An authorship team publishing an update in a scientific or medical journal is likely to manage the new edition of a review in the same way as with any other publication, and follow the ICMJE authorship criteria.[56] If the previous author or author team steps down, then they should be acknowledged in the new version. However, some might perceive that their efforts in the first version warrant continued authorship, which may be valid. The management of authorship between versions can sometimes be complicated. At worst, it delays new authors completing an update and leads to long authorship lists of people from previous versions who probably do not meet ICMJE authorship criteria. One approach with updates including new authors is to have an opt-in policy for the existing authors: they can opt in to the new edition, provided that they make clear their contribution, and this is then agreed with the entire author team.

Although they are new publications, updates will generally include content from the published version. Changing licensing rights around systematic reviews to allow new authors of future updates to remix, tweak, or build on the contributions of the original authors of the published version (similar to the rights available via a Creative Commons licence; https://creativecommons.org) could be a more sustainable and simpler approach. This approach would allow systematic reviews to continue to evolve and build on the work of a range of authors over time, and for contributors to be given credit for contributions to this previous work.

Efficient searching

In performing an update, a search based on the search conducted for the original review is required. The updated search strategy will need to take into account changes in the review question or inclusion criteria, for example, and might be further adjusted based on knowledge of running the original search strategy. The search strategy for an update need not replicate the original search strategy, but could be refined, for example, based on an analysis of the yield of the original search. These new search approaches are currently undergoing formal empirical evaluation, but they may well provide much more efficient search strategies in the future. Some examples of these possible new methods for review updates are described in web appendix 2.

In reporting the search process for the update, investigators must ensure transparency for any previous versions and the current update, and use an adapted flow diagram based on PRISMA reporting (preferred reporting items for systematic reviews and meta-analyses).[57] The search processes and strategies for the update must be adequately reported such that they could be replicated.

Table 2 | Refresh background, objectives, inclusion criteria, and methods

Protocol section	Appraisal points
Background and research question	• Review and update background section, including supporting references to take account of any changes that may have occurred. This should include updating any new information and current policy debates on the topic. • Assess whether the current review question remains relevant to patients and practice.
Inclusion criteria	• Consider whether the existing PICO(s) remain(s) current, in the light of new knowledge. • Identify any new understanding of definition of patient populations. • Identify new interventions, or those that have been withdrawn, that are no longer in use. • Identify any changes in usual care standards. • Check for standardised core outcomes sets, such as those developed in collaboration with the core outcome measures in effectiveness trials (COMET) initiative (www.comet-initiative.org) or by guideline groups since the original review. • Check for any relevant patient reported outcomes to include subsequent to the original review. • Consider any new studies with less risk of bias that might warrant a stricter study design inclusion criteria (where the older version, when there was a dearth of evidence, included observational or quasi-randomised comparisons).
Methods	• Appraise and update the methods pending relevant methodological advancements or developments. For example, if (1) there are new tools for assessing the risk of bias of individual studies or appraising the quality of a body of evidence (eg, GRADE); or (2) new and efficient search approaches are feasible, such as a targeted approach to searching, taking into account the quality of the original search, and ensuring that the search for the update is of high quality. • Update or include a summary of findings table, which is recommended for all systematic reviews, because it improves the clarity, understanding, and interpretation of the findings of a systematic review, and rapidly reduces the amount of time readers require to find key information.[53-55]

ED_004886_00001345-00026

Peer review

Systematic reviews published for the first time in peer reviewed journals are by definition peer reviewed, but practice for updates remains variable, because an update might have few changes (such as an updated search but no new studies found and therefore included) or many changes (such as revise methods and inclusion of several new studies leading to revised conclusions). Therefore, and to use peer reviewers' time most effectively, editors need to consider when to peer review an update and the type of peer reviewer most useful for a particular update (for example, topic specialist, methodologist). The decision to use peer review, and the number and expertise of the peer reviewers could depend on the nature of the update and the extent of any changes to the systematic review as part of an editor assessment. A change in the date of the search only (where no new studies were identified) would not require peer review (except, arguably, peer review of the search), but the addition of studies that lead to a change in conclusions or significant changes to the methods would require peer review. The nature of the peer review could be described within the published article.

Reporting changes

Authors should provide a clear description of the changes in approach or methods between different editions of a review. Also, authors need to report the differences in findings between the original and updated edition to help users decide how to use the new edition. The approach or format used to present the differences in findings might vary with the target user group.[58] Publishers need to ensure that all previous versions of the review remain publically accessible.

Updates can range from small adjustments to reviews being completely rewritten, and the PUGs panel spent some time debating whether the term "new edition" would be a better description than "update." However, the word "update" is now in common parlance and changing the term, the panel judged, could cause confusion. However, the debate does illustrate that an update could represent a review that asks a similar question but has been completely revised.

Technology and innovation

The updating of systematic review is generally done manually and is time consuming. There are opportunities to make better use of technology to streamline the updating process and improve efficiency (table 3). Some of these tools already exist and are in development or in early use, and some are commercially available or freely available. The AHRQ's evidence based practice centre team has recently published tools for searching and screening, and will provide an assessment of the use, reliability, and availability of these tools.[63]

Other developments, such as targeted updates that are performed rapidly and focus on updating only key components of a review, could provide different approaches to updating in the future and are being piloted and evaluated.[64] With implementation of these various innovations, the longer term goal is for "living" systematic reviews, which identify and incorporate information rapidly as it evolves over time.[60]

Concluding remarks

Updating systematic reviews, rather than addressing the same question with a fresh protocol, is generally

Table 3 | Technological innovations to improve the efficiency of updating systematic reviews

Innovation	Description	Application	Examples of software and projects,* and current status
Integrated software	Integration of applying inclusion criteria, review management systems, statistical packages, and GRADE	To facilitate greater efficiencies in review production, including their updates	Covidence (www.covidence.org): free/pay† EPPI reviewer (http://eppi.ioe.ac.uk/cms/Default.aspx): pay DistillerSR (https://distillercer.com/products/distillersr-systematic-review-software/): pay Cochrane Review Manager (http://tech.cochrane.org/revman): free/pay† GRADEpro GDT (http://gradepro.org/): free Rayyan (http://rayyan.qcri.org/): free
Systematic review data repositories	Repositories store information from review (eg, data abstraction forms and the evidence tables)	Improve updating efficiency for new or existing teams as the data abstraction forms, evidence tables, and populating data from the original review are available	Agency of Health Care Research and Quality systematic review data repository (http://srdr.ahrq.gov): operational GRADE database of evidence profiles and evidence to decision frameworks (http://dbep.gradepro.org/): operational
Semi-automation	Machine learning techniques to use alongside human efforts	Finding studies and extracting data could benefit from semi-automation creating time efficiencies[59]	RobotReviewer (http://vortext.systems/robotreviewer): free Cochrane project transform—identification of studies (http://community.cochrane.org/tools/project-coordination-and-support/transform)
Crowdsourcing	Use of volunteers to assist systematic review authors with discrete tasks	Individuals from the "crowd" assist with tasks (identifying and screening studies, translating articles, data extraction) to help in new review production and updates[60-62]	Cochrane Project Transform—crowdsourcing (link as above)
Publication linkage	Ability to link trial registration, trial publications, and reviews citing them will help transparency	This initiative could help identify studies for systematic reviews and could also show the relation between systematic review updates	A cross publisher initiative, CrossRef, is coordinating a threaded publications/linked clinical trial reports initiative to link a clinical trial report (with a trial registration number) report and derivative publications, including reviews (www.crossref.org): operational
Data linkage	Increase links between data, existing software, and reviews	To improve identification and reuse of data for review production and dissemination	Cochrane (http://linkeddata.cochrane.org/): proof of principle example at production phase, but mostly linkage projects at exploratory phase

*Further information can be located on the SR Toolbox site (http://systematicreviewtools.com/).
†Free to Cochrane contributors; other users pay.

ED_004886_00001345-00027

more efficient and allows incremental improvement over time. Mechanical rules appear unworkable, but there is no clear unified approach on when to update, and how implement this. This PUGs panel of authors, editors, statisticians, information specialists, other methodologists, and guideline developers brought together current thinking and experience in this area to provide guidance.

Decisions about whether and when to update a systematic review are judgments made at a point in time. They depend on the currency of the question asked, the need for updating to maintain credibility, the availability of new evidence, and whether new research or new methods will affect the findings.

Whether the review uses current methodological standards is important in deciding if the update will influence the review findings, quality, reliability, or credibility sufficiently to justify the effort in updating it. Those updating systematic reviews to author clinical practice guidelines might consider the influence of new study results in potentially overturning the conclusions of an existing review. Yet, even in cases where new study findings do not change the primary outcome measure, new studies can carry important information about subgroup effects, duration of treatment effects, and other relevant clinical information, enhancing the currency and breadth of review results.

An update requires appraisal and revision of the background, question, inclusion criteria, and methods of the existing review and the existing certainty in the evidence. In particular, methods might need to be updated, and search strategies reconsidered. Authors of updates need to consider inputs to the current edition, and follow ICMJE criteria regarding authorship.[56]

The PUGs panel proposed a decision framework (fig 1), with terms and categories for reporting the decisions made for updating procedures for adoption by Cochrane and other stakeholders. This framework includes journals publishing systematic review updates and independent authors considering updates of existing published reviews. The panel developed a checklist to help judgements about when and how to update.

The current emphasis of authors, guideline developers, Cochrane, and consequently this guidance has been on effects reviews. The checklists and guidance here still applies to other types of systematic reviews, such as those on diagnostic test accuracy, and this guidance will need adapting. Accumulative experience and methods development in reviews other than those of effects are likely to help refine guidance in the future.

This guidance could help groups identify and prioritise reviews for updating and hence use their finite resources to greatest effect. Software innovation and new management systems are being developed and in early use to help streamline review updates in the coming years.

AUTHOR AFFILIATIONS

[1]Cochrane Infectious Diseases Group, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK
[2]Oxford Clinical Trials Research Unit, University of Oxford, Oxford, UK
[3]Cochrane Editorial Unit, Cochrane Central Executive, London, UK
[4]Department of Clinical Epidemiology and Biostatistics and Department of Medicine, McMaster University, Hamilton, ON, Canada
[5]Cochrane GRADEing Methods Group, Ottawa, ON, Canada
[6]Department of Internal Medicine, American University of Beirut, Beirut, Lebanon
[7]Department of Mathematics and Statistics, McMaster University
[8]Evidence-based Practice Center Program, Agency for Healthcare and Research Quality, Rockville, MD, USA
[9]Centre for Reviews and Dissemination, University of York, York, UK
[10]Cochrane Upper Gastrointestinal and Pancreatic Diseases Group, Hamilton, ON, Canada
[11]Lefebvre Associates, Oxford, UK
[12]Kaiser Permanente National Guideline Program, Portland, OR, USA
[13]Iberoamerican Cochrane Centre, Barcelona, Spain
[14]Cochrane Informatics and Knowledge Management, Cochrane Central Executive, Freiburg, Germany
[15]Plymouth University Peninsula School of Dentistry, Plymouth, UK
[16]Department of Clinical Policy, American College of Physicians,Philadelphia, PA, USA
[17]Guidelines International Network, Pitlochry, UK
[18]Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
[19]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[20]Biostatistics Unit, Centre for Evaluation, McMaster University, Hamilton, ON, Canada
[21]Centre for Statistics in Medicine, University of Oxford, Oxford, UK
[22]University of Ottawa, Ottawa, ON, Canada
[23]Cochrane Airways Group, Population Health Research Institute, St George's, University of London, London, UK
[24]Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, UK

Contributors: HJS initiated the workshop. JC, SH, PG, HM, and HJS organised the materials and the agenda. SH wrote up the proceedings. PG wrote the paper from the proceedings and coordinated the development of the final guidance; JC, SH, HM, and HJS were active in the finalising of the guidance. All PUGs authors contributed to three rounds of manuscript revision.

Funding: Attendance at this meeting, for those attendees not directly employed by Cochrane, was not funded by Cochrane beyond the reimbursement of out of pocket expenses for those attendees for whom this was appropriate. Expenses were not reimbursed for US federal government attendees, in line with US government policy. Statements in the manuscript should not be construed as endorsement by the US Agency for Healthcare Research and Quality or the US Department of Health and Human Services.

Competing interests: All participants have a direct or indirect interest in systematic reviews and updating as part of their job or academic career. Most participants contribute to Cochrane, whose mission includes a commitment to the updating of its systematic review portfolio. JC, HM, RM, CM, KS-W, and MT are, or were at that time, employed by the Cochrane Central Executive.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 3.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/3.0/.

1 Shekelle PG, Ortiz E, Rhodes S, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated?*JAMA* 2001;286:1461-7. doi:10.1001/jama.286.12.1461.
2 Claxton K, Cohen JT, Neumann PJ. When is evidence sufficient?*Health Aff (Millwood)* 2005;24:93-101. doi:10.1377/hlthaff.24.1.93.
3 Fenwick E, Claxton K, Sculpher M, et al. *Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Paper 179.*Centre for Health Economics, University of York, 2000.
4 Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty—when is there sufficient evidence?*Value Health* 2005;8:433-46. doi:10.1111/j.1524-4733.2005.00033.x.

ED_004886_00001345-00028

5 Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *J Health Serv Res Policy* 1997;2:26-30.
6 Wilson E, Abrams K. From evidence based economics to economics based evidence: using systematic review to inform the design of future research. In: Shemilt I, Mugford M, Vale L, et al. eds. *Evidence based economics*.Blackwell Publishing, 2010doi:10.1002/9781444320398.ch12.
7 Chalmers I, Enkin M, Keirse MJ. Preparing and updating systematic reviews of randomized controlled trials of health care. *Milbank Q* 1993;71:411-37. doi:10.2307/3350409.
8 Higgins J, Green S, Scholten R. Chapter 3. *Maintaining reviews: updates, amendments and feedback: Version 5.1.0 (updated March 2011)*.Cochrane Collaboration, 2011.
9 Cochrane. Editorial and publishing policy resource. http://community.cochrane.org/editorial-and-publishing-policy-resource. 2016.
10 Moher D, Tsertsvadze A. Systematic reviews: when is an update an update?*Lancet* 2006;367:881-3. doi:10.1016/S0140-6736(06)68358-X.
11 Prasad K, Singh MB, Ryan H. Corticosteroids for managing tuberculous meningitis. *Cochrane Database Syst Rev* 2016;4:CD002244.
12 Zani B, Gathu M, Donegan S, Olliaro PL, Sinclair D. Dihydroartemisinin-piperaquine for treating uncomplicated *Plasmodium falciparum* malaria. *Cochrane Database Syst Rev* 2014;1:CD010927.
13 Adams SP, Tsang M, Wright JM. Lipid lowering efficacy of atorvastatin. *Cochrane Database Syst Rev* 2012;12:CD008226.
14 Higgins J. Convincing evidence from controlled and uncontrolled studies on the lipid-lowering effect of a statin. *Cochrane Database Syst Rev* 2012;12:ED000049.
15 Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ* 2013;347:f7191. doi:10.1136/bmj.f7191.
16 MacLehose H, Hilton J, Tovey D, et al. The Cochrane Library: revolution or evolution? Shaping the future of Cochrane content. Background paper for The Cochrane Collaboration's Strategic Session Paris, France, 18 April 2012. http://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/2012-CC-strategic-session_full-report.pdf.
17 Hemilä H, Chalker E. Vitamin C for preventing and treating the common cold. *Cochrane Database Syst Rev* 2013;(1):CD000980.
18 Hahn S, Kim S, Garner P. Reduced osmolarity oral rehydration solution for treating dehydration caused by acute diarrhoea in children. *Cochrane Database Syst Rev* 2002;(1):CD002847.
19 World Health Organization (WHO). Reduced osmolarity oral rehydration salts (ORS) formulation. A report from a meeting of Experts jointly organized by UNICEF and WHO. New York: Child and Adolescent Health and Development, 18 July 2001 http://apps.who.int/iris/bitstream/10665/67322/1/WHO_FCH_CAH_01.22.pdf.
20 Smaill FM, Grivell RM. Antibiotic prophylaxis versus no prophylaxis for preventing infection after cesarean section. *Cochrane Database Syst Rev* 2014;(10):CD007482.
21 Adjuik M, Babiker A, Garner P, Olliaro P, Taylor W, White N. International Artemisinin Study Group. Artesunate combinations for treatment of malaria: meta-analysis. *Lancet* 2004;363:9-17. doi:10.1016/S0140-6736(03)15162-8.
22 Agnihotry A, Fedorowicz Z, Nasser M. Adhesively bonded versus non-bonded amalgam restorations for dental caries. *Cochrane Database Syst Rev* 2016;3:CD007517.
23 Higgins JP, Altman DG, Gøtzsche PC, et al. Cochrane Bias Methods Group Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928.
24 Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36. doi:10.7326/0003-4819-155-8-201110180-00009.
25 Sterne JAC, Higgins JPT, Reeves BC; on behalf of the development group for ROBINS-I. A tool for assessing risk of bias in non-randomized studies of interventions, version 7. March 2016. www.riskofbias.info.
26 Guyatt GH, Oxman AD, Vist GE, et al. GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD.
27 Schünemann HJ. Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision?*J Clin Epidemiol* 2016;75:6-15. doi:10.1016/j.jclinepi.2016.03.018.
28 Gough D. Qualitative and mixed methods in systematic reviews. *Syst Rev* 2015;4:181. doi:10.1186/s13643-015-0151-y.
29 Richardson M, Garner P, Donegan S. Cluster randomised trials in Cochrane reviews: evaluation of methodological and reporting practice. *PLoS One* 2016;11:e0151818. doi:10.1371/journal.pone.0151818.
30 Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10. doi:10.1186/1471-2288-7-10.
31 Whiting P, Savović J, Higgins JP, et al. ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005.
32 Sampson M, Shojania KG, McGowan J, et al. Surveillance search techniques identified the need to update systematic reviews. *J Clin Epidemiol* 2008;61:755-62. doi:10.1016/j.jclinepi.2007.10.003.
33 Hemens BJ, Haynes RB. McMaster Premium LiteratUre Service (PLUS) performed well for identifying new studies for updated Cochrane reviews. *J Clin Epidemiol* 2012;65:62-72.e1. doi:10.1016/j.jclinepi.2011.02.010.
34 Sagliocca L, De Masi S, Ferrigno L, Mele A, Traversa G. A pragmatic strategy for the review of clinical evidence. *J Eval Clin Pract* 2013;19:689-96. doi:10.1111/jep.12020.
35 Rada G, Peña J, Capurro D, et al. How to create a matrix of evidence in Epistemonikos. Abstracts of the 22nd Cochrane Colloquium; Evidence-informed public health: opportunities and challenges; Hyderabad, India. *Cochrane Database Syst Rev* 2014;suppl 1:132.
36 Okebe JU, Yahav D, Shbita R, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2011;(10):CD006589.
37 Neuberger A, Okebe J, Yahav D, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2016;2:CD006589.
38 Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401-6. doi:10.1016/j.jclinepi.2010.07.015.
39 Chung M, Newberry SJ, Ansari MT, et al. Two methods provide similar signals for the need to update systematic reviews. *J Clin Epidemiol* 2012;65:660-8. doi:10.1016/j.jclinepi.2011.12.004.
40 Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med* 2007;147:224-33. doi:10.7326/0003-4819-147-4-200708210-00179.
41 Shojania K, Sampson M, Ansari M, et al. *Updating systematic reviews; AHRQ technical reviews; report no 07-0087*.Agency for Healthcare Research and Quality, 2007.
42 Pattanittum P, Laopaiboon M, Moher D, Lumbiganon P, Ngamjarus C. A comparison of statistical methods for identifying out-of-date systematic reviews. *PLoS One* 2012;7:e48894. doi:10.1371/journal.pone.0048894.
43 Shekelle PG, Motala A, Johnsen B, Newberry SJ. Assessment of a method to detect signals for updating systematic reviews. *Syst Rev* 2014;3:13. doi:10.1186/2046-4053-3-13.
44 Shekelle PG, Newberry SJ, Wu H, et al. *Identifying signals for updating systematic reviews: a comparison of two methods; report no 11-EHC042-EF*.Agency for Healthcare Research and Quality, 2011.
45 Shekelle P, Newberry S, Maglione M, et al. *Assessment of the need to update comparative effectiveness reviews: report of an initial rapid program assessment (2005-2009)*.Agency for Healthcare Research and Quality, 2009.
46 Tovey D, Marshall R, Bazian, Hopewell S, Rader T. Fit for purpose: centralised updating support for high-priority Cochrane Reviews; National Institute for Health Research Cochrane-National Health Service Engagement Award Scheme, July 2011. https://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/10_4000_01%20Fit%20for%20purpose%20-%20centralised%20updating%20support%20for%20high%20priority%20Cochrane%20Reviews%20FINAL%20REPORT.pdf.
47 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341-64. doi:10.1016/S0167-6296(98)00039-3.
48 Wilson EC. A practical guide to value of information analysis. *Pharmacoeconomics* 2015;33:105-21. doi:10.1007/s40273-014-0219-x.
49 Althabe F, Belizán JM, McClure EM, et al. A population-based, multifaceted strategy to implement antenatal corticosteroid treatment versus standard care for the reduction of neonatal mortality due to preterm birth in low-income and middle-income countries: the ACT cluster-randomised trial. *Lancet* 2015;385:629-39. doi:10.1016/S0140-6736(14)61651-2.
50 Taylor-Robinson D, Maayan N, Soares-Weiser K, et al. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin and school performance. *Cochrane Database Syst Rev* 2012;(11):CD000371.
51 van Dalen EC, van der Pal HJ, Kremer LC. Different dosage schedules for reducing cardiotoxicity in people with cancer receiving anthracycline chemotherapy. *Cochrane Database Syst Rev* 2016;3:CD005008.
52 Wilson E. on behalf of the Cochrane Priority Setting and Campbell & Cochrane Economics Methods Groups. Which study when? Proof of concept of a proposed automated tool to help decision which reviews to update first. *Cochrane Database Syst Rev* 2014;suppl 2:29-31.
53 Rosenbaum SE, Glenton C, Nylund HK, Oxman AD. User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews. *J Clin Epidemiol* 2010;63:607-19. doi:10.1016/j.jclinepi.2009.12.013.
54 Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. *J Clin Epidemiol* 2010;63:620-6. doi:10.1016/j.jclinepi.2009.12.014.

ED_004886_00001345-00029

55 Vandvik PO, Santesso N, Akl EA, et al. Formatting modifications in GRADE evidence profiles improved guideline panelists comprehension and accessibility to information. A randomized trial. *J Clin Epidemiol* 2012;65:748-55. doi:10.1016/j.jclinepi.2011.11.013.
56 International Committee of Medical Journal Editors (ICMJE). Defining the role of authors and contributors. 2016. www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.
57 Stovold E, Beecher D, Foxlee R, Noel-Storr A. Study flow diagrams in Cochrane systematic review updates: an adapted PRISMA flow diagram. *Syst Rev* 2014;3:54. doi:10.1186/2046-4053-3-54.
58 Newberry SJ, Shekelle PG, Vaiana M, et al. *Reporting the findings of updated systematic reviews of comparative effectiveness: how do users want to view new information? report no 13-EHC093-EF* Agency for Healthcare Research and Quality, 2013.
59 Marshall IJ, Kuiper J, Wallace BC. Automating risk of bias assessment for clinical trials BCB'14. Proceedings of the 5th ACM conference on Bioinformatics, computational biology, and health informatics. 2014:88-95. http://thirdworld.nl/automating-risk-of-bias-assessment-for-clinical-trials.
60 Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014;11:e1001603. doi:10.1371/journal.pmed.1001603.
61 Elliott J, Sim I, Thomas J, et al. #CochraneTech: technology and the future of systematic reviews. *Cochrane Database Syst Rev* 2014;(9):ED000091.
62 Cochrane. Project transform: the Cochrane Collaboration. 2016. http://community.cochrane.org/tools/project-coordination-and-support/transform.
63 Paynter R, Bañez L, Berlinerm E, et al. EPC methods: an exploration of the use of text-mining software in systematic reviews. Research white paper. AHRQ publication 16-EHC023-EF. Agency for Healthcare Research and Quality, April 2016. https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2214.
64 Soares-Weiser K, Marshall R, Bergman H, et al. Updating Cochrane Reviews: results of the first pilot of a focused update. *Cochrane Database Syst Rev* 2014;suppl 1:31-3.

© BMJ Publishing Group Ltd 2016

Web appendix: Supplementary material

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

September 19, 2017

Comments from Academics, Scientists and Clinicians on the Risk Evaluation Scoping Efforts Under TSCA for Ten Chemical Substances

Submitted online via *Regulations.gov* to dockets: EPA-HQ-OPPT-2016-0736, EPA-HQ-OPPT-2016-0741, EPA-HQ-OPPT-2016-0733, EPA-HQ-OPPT-2016-0723, EPA-HQ-OPPT-2016-0735, EPA-HQ-OPPT-2016-0742, EPA-HQ-OPPT-2016-0743, EPA-HQ-OPPT-2016-0732, EPA-HQ-OPPT-2016-0725, EPA-HQ-OPPT-2016-0737

These comments are submitted on behalf of the undersigned academic, scientists, and clinicians. We declare collectively that we have no direct or indirect financial or fiduciary interest in any chemical under consideration in these risk evaluations. The co-signers' institutional affiliations are included for identification purposes only and do not necessarily imply any institutional endorsement or support, unless indicated otherwise.

We appreciate the opportunity to provide written comments on the scope of risk evaluations for the first ten chemical substances for risk evaluations pursuant to the Toxic Substances Control Act (TSCA), as amended by the Frank R. Lautenberg Chemical Safety of the 21st Century Act (Lautenberg TSCA). Collectively, these chemicals represent an aggregate production volume of more than 1 billion pounds a year in 2015.[1] Some of these chemicals have assessments, and in some cases even restrictions, under other federal programs – but none of these other programs has the mandate given to EPA under the new TSCA: to comprehensively evaluate chemicals and ensure that they do not pose an unreasonable risk to human health and the environment, with special consideration to those most vulnerable amongst us. Therefore, the task ahead for EPA is critical.

These first ten evaluations are also consequential because they will be precedent setting for the implementation of evaluation of science under TSCA. The consequent health impacts of EPA's decisions – for better or worse – will be borne by generations of American children, workers, families, and communities. With so much at stake, we welcome EPA's engagement with the public in this process and we offer EPA concrete approaches to embed the most current scientific principles in its methods to assess the hazards and risks of environmental chemicals.

Our comments address the following main points:

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**
2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**
3. **EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations**

[1] This is the aggregate production volume estimate for the 8 chemicals with production volume information available. For asbestos and pigment violet 29, manufacturers/ importers claimed production volumes as confidential business information (CBI).

ED_004886_00001477-00001

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**
5. **For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.**
6. **Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.**

We are appreciative of the opportunity to provide public input and we look forward to continuing to participate in such opportunities in the near future. Please do not hesitate to contact us with any questions regarding these comments.

Sincerely,

Veena Singla, PhD; Juleen Lam, PhD; and Tracey Woodruff, PhD
Program on Reproductive Health and the Environment
University of California, San Francisco

Patricia D. Koman, MPP, PhD
President and Senior Health Scientist
Green Barn Research

Kathy Attar, MPH
Toxics Program Manager
Physicians for Social Responsibility

Sheila Brear, BDS
Associate Dean of Academic Affairs, School of Dentistry
University of California, San Francisco

Phil Brown, PhD
University Distinguished Professor of Sociology and Health Sciences
Northeastern University

Adelita G. Cantu, PhD, RN
Associate Professor
Alliance of Nurses for Healthy Environments

Courtney Carignan, PhD
Assistant Professor
Department of Food Science and Human Nutrition and Department of Pharmacology and Toxicology
Michigan State University

ED_004886_00001477-00002

Jeanne A. Conry, MD, PhD
President, Environmental Health Leadership Foundation
Past President, The American Congress of Obstetricians and Gynecologists

Carl F. Cranor, PhD
Distinguished Professor of Philosophy
Faculty Member, Environmental Toxicology Graduate Program
Department of Philosophy
University of California, Riverside

Adam M. Finkel, ScD, CIH
Clinical Professor of Environmental Health Sciences, University of Michigan School of Public Health
Executive Director, Penn Program on Regulation, University of Pennsylvania Law School

Susan J. Fisher, PhD
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Robert M. Gould, MD
Associate Adjunct Professor
Program on Reproductive Health and the Environment
University of California, San Francisco
Past President, Physicians for Social Responsibility

Alycia Halladay, PhD
Chief Science Officer
Autism Science Foundation

Jean-Marie Kauth, PhD
Associate Professor
Benedictine University

Detlef R.U. Knappe, PhD
Professor
Department of Civil, Construction, and Environmental Engineering
North Carolina State University

Erica Koustas, PhD
Scientific Consultant
Program on Reproductive Health and the Environment
University of California, San Francisco

Hal C. Lawrence, III, MD, FACOG*
Executive Vice President and CEO
American Congress of Obstetricians and Gynecologists

ED_004886_00001477-00003

Heather Patisaul, PhD
Professor, Biological Sciences
Center for Human Health and the Environment, WM Keck Center for Behavioral Biology
North Carolina State University

Joshua F. Robinson, PhD
Assistant Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Ted Schettler MD, MPH
Science Director
Science and Environmental Health Network

Patrice Sutton, MPH
Research Scientist
Program on Reproductive Health and the Environment
University of California, San Francisco

Aolin Wang, PhD
Postdoctoral Scholar
Program on Reproductive Health and the Environment
University of California, San Francisco

Nsedu Obot Witherspoon, MPH
Executive Director
Children's Environmental Health Network

Lauren Zajac, MD, MPH
Assistant Professor
Department of Environmental Medicine and Public Health
Icahn School of Medicine at Mount Sinai

Marya G. Zlatnik, MD, MMS
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

*indicates organizational support

ED_004886_00001477-00004

DETAILED COMMENTS

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**

Overall, we strongly commend the EPA for its efforts to utilize a systematic and transparent method of research synthesis to reach a concise, strength of evidence conclusion about the human health hazard resulting from exposures to these ten chemicals. Efforts to integrate systematic review methods, including the explicit development of search terms, strategies, and inclusion/exclusion criteria beforehand, is relatively new in EPA's chemical assessment and as such, we applaud the EPA for this and its general improvements in its hazard assessment methodology. These scoping documents generally provide an important infrastructure for outlining EPA's screening approach for identifying relevant references and to document decisions made in the process of identifying the body of scientific literature that will be evaluated in the chemical assessments.

To improve on this document and advance EPA's uptake of systematic review methods of research synthesis, we identify the following opportunities for improvement:

EPA should not exclude studies based on language. EPA's search strategy is limited to English-only studies. The exclusive reliance on English-language studies may not represent the entire body of available evidence, and studies have suggested that language bias might lead to erroneous conclusions.[2] Furthermore, when considering the inclusion or update of an existing systematic review, studies have found that language-inclusive systematic reviews (including studies in languages other than English) were of the highest quality, compared with other types of reviews.[3] Online translation tools are readily available to allow screeners to quickly evaluate study abstracts for relevance, and therefore we recommend EPA to incorporate non-English language studies in their screening and not simply exclude these potentially relevant papers.

EPA should provide exclusion reasons for off topic citations. In the Bibliography Supplemental File for the Scope Documents, EPA has provided lists of bibliographic citations that were identified and screened from the initial literature search and the initial categorization of whether citations were *on topic* or *off topic.* We recommend EPA additionally provide exclusion reasons that were used to come to the conclusion that each citation was *off topic*, as this is a standard recommendation to fulfill transparency in documenting and reporting all decisions made in the study selection process.[4] This is particularly important as EPA has proposed to do its screening in Distiller, proprietary software that presumably will

---

[2] Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., Mierzwinski-Urban, M., Clifford, T., Hutton, B. and Rabb, D., 2012. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *International journal of technology assessment in health care*, *28*(2), pp.138-144.

[3] Moher, D., Pham, B., Lawson, M.L. and Klassen, T.P., 2003. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess*, *7*(41), pp.1-90.

[4] McDonagh M et a. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Agency for Healthcare Research and Quality. 2008.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J. and Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, *6*(7), p.e1000100.

Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. March 2011. Available from: http://handbook-5-1.cochrane.org/chapter_7/7_2_5_selecting_excluded_studies.htm

ED_004886_00001477-00005

not be made publicly available, raising concerns regarding the transparency and reproducibility of this screening step.

EPA should consider other tools for systematic review. EPA has also proposed to extract data results in the DRAGON software. We strongly encourage EPA to also consider other potential software tools that have been developed and actively incorporated into the process of systematic review, such as Swift Reviewer,[5] Active Screener,[6] HAWC (Health Assessment Workplace Collaborative).[7] These tools will help to ensure consistent and transparent execution and presentation of reviews and increase transparency of EPA assessment. Furthermore, we urge EPA to work with the National Toxicology Program and other organizations involved in these efforts in an ongoing basis to develop these and other open source tools to train scientists in their use. We believe that such infrastructure development will be critical to increasing the efficiency of chemical assessments and to expedite uptake of systematic reviews in environmental health.

EPA should have two independent reviewers for screening steps. EPA has outlined its process for screening title and abstracts of papers as having a single reviewer reviewing papers to determine whether the study is *on-topic* or *off-topic*. As part of this process, a senior-level technical expert in the topic area of interest reviewed the appropriateness of the assigned tags for "the first batch of studies" and provided feedback to the screener. Senior-level technicians also provided feedback and guidance on specific references to the individual screeners as needed during the screening and tagging process. From the description of this process, it does not appear that two independent reviewers screen all titles and abstracts for potential inclusion. Using two independent reviewers is a standard approach in systematic reviews and therefore we strongly recommend that EPA include a second independent reviewer within this process to ensure that **all** studies are screened by two reviewers at each step (title and abstract and full text).

EPA should clearly document decisions related to the identification and search. For example, it was unclear how many studies were included in the first batch of studies reviewed by the senior-level technician—these decisions should be clearly specified beforehand as to the number (or percent) that will be reviewed by this independent reviewer. Furthermore, it is unclear how many studies the senior-level technical experts are reviewing generally as to their additional feedback and guidance to individual screeners. This should be more clearly stated and described beforehand in these protocols.
We recommend EPA broaden the set of studies that are initially screened in the first batch to ensure consistency across reviewers and demonstrated understanding of protocol instructions by all reviewers before moving on to screening the remaining records. It is stated in the Gray Literature Search Results that individual screeners would screen and tag 10 references that would be then independently reviewed by the senior-level technical expert. However, this does not seem to be an adequate number of studies as it is a small number relative to the expected number of records that will ultimately be screened.

EPA should clarify how it will handle discrepancies in the inclusion/exclusion and tagging process. As it is stated in its current protocol, it appears that the senior-level technical expert has the final say in

---

[5] Howard, B.E., Phillips, J., Miller, K., Tandon, A., Mav, D., Shah, M.R., Holmgren, S., Pelch, K.E., Walker, V., Rooney, A.A. and Macleod, M., 2016. SWIFT-Review: a text-mining workbench for systematic review. *Systematic reviews*, *5*(1), p.87.
[6] https://www.sciome.com/swift-activescreener/
[7] https://hawcproject.org/about/

determining the final inclusion/exclusion decision and tagging, for the subset of studies they evaluate. However, this should be clarified and we also highly recommend that a third party reviewer be incorporated as an arbiter for these decisions if consensus cannot be reached between the two reviewers, as is typically standard in systematic reviews.

EPA should stratify its exclusion criteria separately at the title and abstract and full-text screening steps. It is likely that title and abstracts of papers would not contain sufficient detail to evaluate all exclusion criteria—many of these would likely only be identified in the full-text of the paper. To increase the efficiency of the screening process, it would help to create a subset of exclusion criteria most relevant when screening the title and abstracts of records versus the larger set of exclusion criteria relevant to screening the full text of records.

EPA should clearly outline the process for handling anticipated overlap with literature relevant to multiple topics. EPA should describe how this will be addressed by the screeners and whether the same reviewer will be responsible for screening papers with inclusion/exclusion criteria across multiple topics or whether different reviewers are responsible only for screening studies for one particular topic. Additional details in regards to the process by which this screening will occur would be helpful. Given the breadth of each assessment (searching literature related to fate, engineering, exposure, human health, and environmental hazard) and the complexity of the screening process (tagging *on-topic* and *off-topic* literature and using additional sub-categories or sub-tags to allow for additional categorization), there appears to be the potential for individual papers to fall into different topic categories and have many different tags and sub-tags applied to indicate their relevance. However, it is unclear how this will be organized in the screening phase. Search strategies and inclusion/exclusion criteria appear to have been developed specifically for each literature topic and the potential overlap of literature relevant to multiple topics is not addressed.

EPA should explicitly include stopping rules for when the list of relevant studies will be considered final. There is no discussion of stopping dates or the process of updating the literature search to search for newer studies. Newer scientific studies will inevitably continue to appear in scientific journals and it will be impossible to continually attempt to include all these studies in a chemical assessment. To meet the deadlines as mandated by the Lautenberg Amendments, EPA should state clear stopping rules in the form of deadlines or criteria for when the body of included relevant studies will be finalized for the purposes of the chemicals assessment.

EPA should ensure gray literature search results are adequately screened. EPA's gray literature search strategy proposes to utilize Google's API to develop custom searches and return the first 100 results, sorted by predicted relevancy so that the results likely to be most relevant are screened first. It is unclear why this number is limited to only the first 100 and whether there was an empirical reason for why this particular number was selected. We recommend that EPA ensure that an adequate number of search results are screened, in particular considering that the gray literature can contribute potentially important information relevant to toxicity, mode of action, exposure, fate and transport, engineering or occupational exposure, or existence of publication bias.

EPA should consider "snowball searching," where the citations of included (i.e., *on-topic*) references are searched as well as using databases such as Web of Science to search for references that cite the included citations. EPA states that it plans on assessing the specificity and efficiency of the literature searches, through comparison of references either cited in existing problem formulation and risk assessment documents, in the public use documents and supporting life cycle diagrams, and comparison

of the references cited in review articles. Snowball searching will contribute to the evaluation of the specificity and efficiency of its literature searches, and also help to identify newer relevant studies that could potentially be included that have not yet been indexed in main databases such as PubMed.

EPA should incorporate appropriate tools for updating and evaluating systematic reviews in their chemical assessments. Garner et al., as part of efforts by a Cochrane Collaboration panel for updating guidance for systematic reviews, published guidance in 2016 for determining when it is appropriate to update a systematic review and outlining the steps for performing the update.[8] We have attached this guidance as an Appendix to these comments. EPA should evaluate the Cochrane tool's applicability to environmental chemicals given that Cochrane systematic reviews are geared towards reviews of clinical intervention evidence, so these tools may require updating and tailoring for an application to environmental health data.

A recent NAS report[9] recommends EPA should develop policies and procedures to allow the agency to identify, use and update existing systematic reviews. The committee also noted that it was important that the existing systematic review's study question directly addresses EPA's topic of interest and that the methods are critically evaluated before the systematic review is used and updated. EPA should ensure that only the highest quality systematic reviews be considered appropriate for use. It will be critical for EPA to develop tools to assist with the process of evaluating existing systematic reviews, particularly as this field continues to rapidly expand and more systematic reviews relevant to environmental health questions are published in the scientific literature, potentially of variable quality.

One tool which might be helpful for evaluating the risk of bias in systematic reviews is the ROBIS tool, which the NAS committee utilized in their report.[10] This tool was developed using rigorous methodology and can be applied for evaluating internal validity of systematic reviews in conjunction with other available tools to critically appraise and assess their quality. Of particular note is the strong emphasis on the recommendation that tools such as ROBIS should not be used to generate a composite quality score, as it has been well-documented that scoring can lead to bias in evaluation of the studies.[11] As such, the ROBIS tool presents several options for visually and graphically presenting results from risk of bias assessments based on individual domains or the overall rating, enabling reviewers to highlight particular areas of concern or reviews that are most relevant to the target question of interest.

Another tool which may be helpful in this process is the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), used by authors of systematic reviews to improve the reporting

---

[8] Garner P, Hopewell S, Chandler J, MacLehose H, Akl EA, Beyene J, Chang S, Churchill R, Dearness K, Guyatt G, Lefebvre C. When and how to update systematic reviews: consensus and checklist. bmj. 2016 Jul 20;354:i3507.

[9] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[10] Id.

[11] Jüni, P., Witschi, A., Bloch, R. and Egger, M., 1999. The hazards of scoring the quality of clinical trials for meta-analysis. *Jama*, *282*(11), pp.1054-1060.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

ED_004886_00001477-00008

of elements relevant to the systematic review and meta-analyses.[12] Increasingly, scientific journals are requiring the inclusion of checklists such as PRISMA with the submission of systematic reviews considered for publication. Although this tool is used to evaluate study reporting, and is not an assessment instrument to gauge the quality of a systematic review, it can still provide a useful framework to identify reported components of an existing systematic review in the process of evaluating quality or to identify missing components requiring follow-up with study authors to obtain additional information.

Furthermore, we strongly encourage EPA to evaluate the potential for financial conflicts of interest as an element in their study design. This is currently included as a consideration in evaluation risks of bias in some frameworks, such as the Navigation Guide,[13] and extracted for consideration as an additional domain in other frameworks, such as that developed by the National Toxicology Program's (NTP) Office of Health Assessment and Translation (OHAT).[14] The Cochrane Collaboration's risk of bias tool does not currently include a specific domain for bias related to study funding source, but this is an area of active discussion among its members.[15] The Cochrane Collaboration has recognized the importance of identifying study funding source, which has been empirically shown to be associated with biases.[16] A recent report from the NRC recommended that the U.S. EPA consider funding sources in their risk of bias assessment conducted for systematic reviews.[17]

We also strongly recommend EPA identify tools that may potentially not be appropriate for human health chemical assessments. Many tools are currently being developed for evaluating risk of bias, quality, and strength of evidence for individual studies as well as for systematic reviews. It is critical that EPA evaluate tools developed in other fields that may be relevant, such as for clinical or preclinical animal or human studies, as these tools could potentially be modified for an application to questions of environmental health relevance. However, these tools should be applied with caution—due to the

---

[12] http://www.prisma-statement.org/; Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews. 2015 Jan 1;4(1):1.

[13] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

[14] NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[15] Bero LA. Why the Cochrane risk of bias tool should include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000075.*

Sterne JA. Why the Cochrane risk of bias tool should not include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000076.*

[16] Krauth D, Anglemyer A, Philipps R, Bero L. Nonindustry-sponsored preclinical studies on statins yield greater efficacy estimates than industry-sponsored studies: a meta-analysis. *PLoS Biol. 2014 Jan; 12(1):e1001770*

Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev. 2012 Dec 12; 12():MR000033.*

[17] NRC (National Research Council). Review of EPA's Integrated Risk Information System (IRIS) Process. Washington, DC:National Academies Press. 2014.

ED_004886_00001477-00009

differences in the types of evidence under evaluation a direct application to a difference evidence base than intended could lead to biased and erroneous conclusions.[18]

2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**

In general, EPA is proposing to consider three populations for exposure assessment: 1) Occupational users and non-users; 2) consumers and bystanders; and 3) general population. We strongly recommend that EPA calculate the aggregate exposures within and across these populations-- risk will be underestimated if it does not include these real-world exposures. Exposures within a population should also be aggregated (rather than considered in isolation) in order to estimate the general population's actual exposure to the chemical—for example, through exposures from food, water and air.

Further, as shown in the Figure below, exposures must also be aggregated across populations. Consumers and workers are part of the general population – that is, since workers and consumers also eat food and drink water, they will have the same exposures as the general population, in addition to the anticipated exposures on-the-job or from consumer products. Some workers will also be consumer product users, so they have the potential to face general, consumer product, and on-the-job exposures. These specific exposure scenarios must be accounted for in EPA's exposure estimation to ensure that such individual exposures are adequately considered and integrated into the risk assessment.

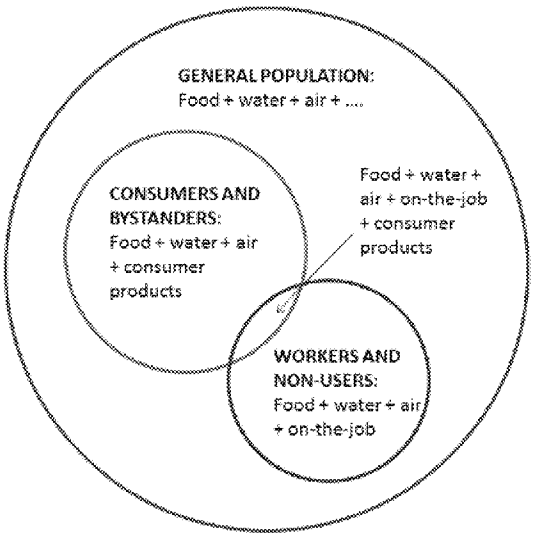


Figure: EPA must assess aggregate exposures within and across all the populations for accurate exposure assessment.

In the Introduction section of the chemical Scope documents, EPA states that it "may consider background exposures from legacy use, associated disposal, and legacy disposal as part of an assessment of aggregate exposure or as a tool to evaluate the risk of exposures resulting from non-

[18] Nachman, K.E., Lam, J., Schinasi, L.H., Smith, T.C., Feingold, B.J. and Casey, J.A., 2017. O'Connor et al. systematic review regarding animal feeding operations and public health: critical flaws may compromise conclusions. *Systematic Reviews*, *6*(1), p.179.

legacy uses."[19] This falls short of the analysis required under Lautenberg TSCA. It is critical that EPA consider ongoing exposures from legacy uses and disposal, and includes these as part of the aggregate exposure assessment. Asbestos and HBCD are two examples of this, as they have enormous volumes in place in buildings and existing infrastructure. The Healthy Building Network estimates there are 66 million- 132 million pounds (30,000-60,000 metric tons) of HBCD in insulation in existing buildings[20] — these reservoirs in-place are and will continue to be critical sources of ongoing exposures. HBCD was also used in cars and furniture, which are long-lived consumer items that will continue to contribute to ongoing exposures for years to come.

Another example is 1,4-dioxane, which was historically used as a chemical stabilizer for chlorinated solvents. Many groundwater aquifers are contaminated with 1,4-dioxane, and the extent of legacy contamination of groundwater is likely underestimated.[21] Also, 1,4-dioxane occurs in a wide variety of products including personal care products, detergents, waxes, and antifreeze, and 1,4-dioxane is a by-product in manufacturing processes involving ethylene oxide, such as the production of polyethylene terephthalate (PET), polyester, and surfactants. The use and disposal of 1,4-dioxane has led to past environmental contamination which contributes to on-going exposures.[22] The physical and chemical properties of 1,4-dioxane render it a persistent and highly mobile water contaminant: it is highly miscible in water.[23] Exposures via drinking water are documented back to the 1980s and continue today.[24] Results from EPA's Third Unregulated Contaminant Monitoring Rule (UCMR3) highlight that over 13% of 4,905 public drinking water systems serving >10,000 people had concentrations of 1,4-dioxane above the EPA Reference Concentration of 0.35 ppb 1,4-dioxane.[25] Furthermore, the UCMR3 results do not capture exposures in communities served by small public drinking water systems serving <10,000 people.[26] Approximately 27% of the US population is served by small public drinking water systems.[27] Thus, it will be critical for EPA to consider the population's current exposure to 1,4-dioxane via sources like drinking water as part of their assessment for health risks.

---

[19] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 12

[20] Safer Chemicals, Healthy Families et al. Comments to the U.S. Environmental Protection Agency (EPA) on the Scope of its Risk Evaluation for the TSCA Work Plan Chemicals: CYCLIC ALIPHATIC BROMIDE CLUSTER or HEXABROMOCYCLODODECANE (HBCD). March 15, 2017. https://healthybuilding.net/uploads/files/saferchemicals-hbcd.pdf

[21] David T. Adamson, Shaily Mahendra, Kenneth L. Walker, Jr., Sharon R. Rauch, Shayak Sengupta, and Charles J. Newell. A Multisite Survey To Identify the Scale of the 1,4-Dioxane Problem at Contaminated Groundwater Sites. *Environmental Science & Technology Letters* 2014 1 (5), 254-258 DOI: 10.1021/ez500092u

[22] Agency for Toxic Substances and Disease Registry. TOXICOLOGICAL PROFILE FOR 1,4-DIOXANE. Atlanta, GA; 2012.

[23] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

[24] Dietrich, A.M., D.S. Millington, and R.F. Christman. 1983. Specific identification of organic pollutants in Haw River water using gas chromatography/mass spectrometry. WRRI Report No. 206. North Carolina State University, Raleigh, NC.

[25] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

U.S. Environmental Protection Agency. Unregulated Contaminant Monitoring Rule 74 FR 51850, EPA-HQ-OW-2007-1189 FRL-8963-6. Washington, D.C.; 2009. https://federalregister.gov/a/E9-24287.s

[26] Knappe, D.R.U., Lopez-Velandia, C., Hopkins, Z., & Sun, M. (2016). Occurrence of 1,4-Dioxane in the Cape Fear River Watershed and Effectiveness of Water Treatment Options for 1,4-Dioxane Control. NC Water Resources Research Institute.

[27] EPA. 2017. Small Drinking Water Systems Research https://www.epa.gov/water-research/small-drinking-water-systems-research-0 (accessed September 17, 2017).

ED_004886_00001477-00011

When a chemical is present in products or media as a contaminant/ by-product, EPA needs to include and assess these exposures. We strongly recommend against ignoring or discounting these potential exposures routes. For example, EPA proposes to exclude from consideration conditions of use of 1,4-dioxane when it is present as contaminant in a wide variety of items, including household detergents, cosmetics/ toiletries, and foods.[28] This exclusion is not scientifically justified. Cosmetics and personal care products have the potential to contribute significantly to exposures, since people are applying them directly to their bodies, often multiple times per day, every day.

Finally, in the exposure assessments for methylene chloride, N-methylpyrrolidone and trichloroethylene, EPA is proposing to exclude uses it already assessed.[29] We agree that EPA does not need to re-assess these uses; these evaluations have been completed and finalized. However, unless and until such uses are banned, the exposures from these uses continue. Therefore, the new risk evaluations need to consider the contributions of these uses to exposures by using the exposure values from the previous assessments.

For the occupational exposure analysis plan, EPA states it will "Consider and incorporate applicable engineering controls and/or personal protective equipment into exposure scenarios."[30] However, these are not realistic assumptions nor are they appropriate for public health protection. EPA's own research shows that the primary factors influencing whether a user understands label information are the users' literacy and numeracy, which frequently correlate with the users' education and income.[31] Therefore, people with less education, lower income, and less advanced literary skills will be the most likely to not understand label instructions. These individuals already disproportionately bear the burden of exposures to multiple environmental hazards and the resulting health impacts; thereby placing further burden on this already stressed susceptible subpopulation. Further, appropriate personal protective equipment (PPE) for workers is often not provided by employers, or may not be fitted or working properly. When evaluating occupational exposures, EPA needs to take into consideration all potential and feasible routes of exposure, and should not exclude exposure routes based on assumptions of PPE and/ or exposure controls in place. These controls are not guaranteed and may change in the future, so to assume zero exposure via these routes would be inappropriate and a failure to adequately ensure health protections, especially for susceptible sub-populations as required by the Lautenberg TSCA.

In summary, EPA needs to account for all the sources of exposure or it will underestimate risk for all 10 chemicals. When analyzing aggregate exposures, "sentinel exposure" may be considered simultaneously, where appropriate. However, these are not mutually exclusive and EPA should not incorporate sentinel to the exclusion of aggregate.

3. EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations

---

[28] US EPA (2017). Scope of the Risk Evaluation for 1,4-dioxane. Pg. 21

[29] US EPA (2017). Scope of the Risk Evaluation for Methylene Chloride (Dichloromethane, DCM). Pg. 30
US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 19-20
US EPA (2017). Scope of the Risk Evaluation for Trichloroethylene. Pg. 27

[30] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 45

[31] US Environmental Protection Agency (EPA). 2016. "The Effectiveness of Labeling on Hazardous Chemicals and Other Products." Office of Chemical Safety and Pollution Prevention. RIN 2079-AK07.

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

In general, EPA proposes to consider workers and occupational non-users, consumer and by-standers, and other groups within the general population in proximity to conditions of use as sub-populations who experience greater exposures. In particular, EPA has appropriately identified people who live or work near manufacturing, processing, distribution, use or disposal sites as facing greater exposures. Such communities are often low income and/ or people of color, exposed to a disproportionate share of pollution, environmental hazards, social and economic stressors. Multiple exposures to chemical and non-chemical stressors collectively increase the risk of harm, combined with synergistic effects with other health stressors in their daily lives such as limited access to quality health care.[32,33]

EPA's risk evaluation needs to fully account for the reality of cumulative exposures, as recommended by the NAS in their Phthalates and Cumulative Risk report.[34] As described below, EPA can use "default values" to account for cumulative exposures.

In regards to greater susceptibility, EPA's considerations for addressing susceptibility vary considerably across the 10 chemicals. EPA should apply a consistent approach to addressing susceptibility across the 10 chemicals. The following are well-known factors that increase biologic sensitivity or reduce resilience to exposures,[35,36] and these as well as other relevant factors should be standard considerations for all 10 chemicals to identify susceptible sub-populations:

Intrinsic/ endogenous factors

- Genetic polymorphisms/ genetics/ genetic makeup
- Health status/ nutritional status/ disease status/ pre-existing conditions
- Prenatal lifestage
- Age

Extrinsic factors

- Multiple exposures/ co-exposures
- Race/ ethnicity
- Socioeconomic status (SES)

For example, the prenatal lifestage is the most sensitive to developmental and reproductive toxicants, and women of child-bearing age should be considered as a susceptible sub-population for any chemicals with such hazards. Women of reproductive age are not specifically identified as a potential susceptible

---

[32] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[33] Vesterinen HM, Morello-Frosch R, Sen S, Zeise L, Woodruff TJ. Cumulative effects of prenatal-exposure to exogenous chemicals and psychosocial stress on fetal growth: Systematic-review of the human and animal evidence. Meliker J, editor. PLoS One. 2017 Jul 12;12(7):e0176331.

[34] National Research Council. Committee on the Health Risks of Phthalates, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies. 2008. Phthalates and cumulative risk assessment: the task ahead. Washington, D.C.: National Academies Press.

[35] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[36] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

sub-population for pigment violet 29, TCE, NMP, PERC, or HBCD, even though EPA will consider reproductive and developmental toxicity hazards for these chemicals.

As discussed below, science-based defaults should be used to account for these and other susceptibilities, unless there is there is chemical-specific data available to support increasing or decreasing the default.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**

EPA cites existing IRIS assessments for five chemicals; because these are EPA's own assessments which have gone through the Agency's peer-review process, and in some cases NAS review, EPA can rely on these existing finalized, authoritative assessments for hazard identification.

Moving forward, a weight of evidence evaluation is required by law, which EPA defines as:

> "Weight of scientific evidence means a systematic review method, applied in a manner suited to the nature of the evidence or decision, that uses a pre-established protocol to comprehensively, objectively, transparently, and consistently, identify and evaluate each stream of evidence, including strengths, limitations, and relevance of each study and to integrate evidence as necessary and appropriate based upon strengths, limitations, and relevance."[37]

Therefore, EPA should use a systematic review process for evaluating scientific information for chemicals that do not have an IRIS assessment and for any additional studies that will be considered for the chemicals that have IRIS assessments.

For the scoping document, EPA should include all hazards identified in the literature, and not make decisions about their relevance to the risk evaluation until a systematic review has been completed. For a number of chemicals, EPA has inappropriately drawn conclusions about hazards prior to the completion of a systematic review. Some examples are given in the table below where EPA concludes that HBCD, NMP and pigment violet 29 are not genotoxic based on previous assessments and without conducting a systematic review.

Chemical	Example Text from EPA Scoping Document
HBCD	"Available data suggest that HBCD is not genotoxic. Existing assessments have also concluded, based on genotoxicity information and a limited lifetime study, that HBCD is not carcinogenic (NICNAS, 2012; EINECS, 2008; TemaNord, 2008; OECD, 2007). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity or cancer hazards in the risk evaluation of HBCD at this time."[38]
NMP	"NMP is not mutagenic, based on results from bacterial and mammalian *in vitro* tests and *in vivo* systems and is not considered to be carcinogenic (RIVM, 2013; OECD,

[37] 82 Fed. Reg. 138, 33748
[38] US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg.36

	2007; WHO, 2001). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity and cancer hazards in the NMP risk evaluation."[39]
Pigment violet 29	"Testing for carcinogenicity of Pigment Violet 29 has not been conducted. However, negative genotoxicity results, structure-activity considerations and the expectation of negligible absorption and uptake of Pigment Violet 29 (based on very low solubility), indicate carcinogenicity of Pigment Violet 29 is unlikely. Unless new information indicates otherwise, EPA does not expect to conduct additional, in-depth analyses of genotoxicity and cancer hazards in the risk evaluation of Pigment Violet 29."[40]

The National Academies recently released a report with recommendations on implementation of systematic review for EPA's chemical evaluations (which we will refer to as the 'NAS Systematic Review report' for simplicity).[41] First, they recommend that EPA should develop policies and procedures that allow the agency to use and update existing systematic reviews, since the committee concluded that could potentially save time and resources. EPA should conduct a review to determine whether there are existing systematic reviews on the topic of interest and if there is, EPA should evaluate it to determine if it is high-quality. The NAS recommends that EPA should build on existing high-quality reviews to incorporate new studies and use the updated systematic review as a basis for its assessment. The assessments cited by EPA to support the hazard identification claims are not systematic reviews; even if they were, EPA should evaluate them for quality before relying on their conclusions.

Second, it is very likely that additional studies have been published since the assessments EPA cites were completed. EPA should develop criteria to evaluate the internal validity (risk of bias) of individual studies, utilizing existing tools that have been developed and empirically demonstrated on environmental health studies such as the Navigation Guide or the OHAT approach.[42] We also recommend that EPA not using a scoring system to evaluate study quality. Specifically, we note that empirically validated approaches in the clinical sciences such as Cochrane discourage using a numerical scale scoring approach for evaluating study quality because calculating a score requires choosing a weighting scheme for each component, which generally is nearly impossible to justify.[43] Furthermore, a study might be well designed to eliminate bias, but because the study failed to report details in the publication under review, it will receive a low score--most available scoring systems include a mix of risk of bias and reporting biases which is inappropriate. Additionally, quality scores have been shown to be invalid for assessing risk of bias in clinical research.[44] The current standard in evaluation of both clinical

[39] US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 36

[40] US EPA (2017). Scope of the Risk Evaluation for Pigment Violet 29. Pg.29

[41] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[42] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[43] Juni, P., A. Witschi, R. Bloch, and M. Egger. 1999. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 282(11):1054-1060.

[44] Id.

ED_004886_00001477-00015

and environmental health research calls for reporting each component of the assessment tool separately and not calculating an overall numeric score.[45]

Data generated by alternative test methods (such as high-throughput screening methods) are not different than any other type of *in vitro* or cell-based assay data that would be considered in a systematic review. These kinds of assays provide mechanistic data, and the NAS Systematic Review report explicitly considered how mechanistic data could be utilized in a systematic review for evidence integration. The committee came to two conclusions. First, the same protocol for evaluating relevance and study quality must be used with mechanistic data as for any other study. For example, in the report's case study on phthalates, the committee was not able to integrate results from high-throughput assays because the cell lines used were of unknown relevance to the *in vivo* mechanism of phthalate toxicity.[46] Second, the foundation of the hazard classification in a systematic review is the animal and human data, with the mechanistic data playing a supporting role. If mechanistic data is relevant, it can be used to upgrade a hazard classification, or increase the confidence of a finding made based on evaluation of animal and human data. A hazard classification is never made based on high-throughput or other kinds of mechanistic data alone.[47]

5. For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.

Defaults

We strongly support the use of health protective defaults to incorporate factors that reflect the range of variability and susceptibility in the population to ensure risks are not underestimated. The importance of using protective science-based defaults was highlighted by the NAS in 2009.[48] The default should be used for factors that are known to influence risk unless there is chemical-specific data that support increasing or decreasing it; when there is inadequate information to quantitatively assess inter- or intra-species differences for a specific chemical, the defaults should be used. For example, EPA's defaults should include:

- Inter-human variability, general
- Inter-human susceptibility to carcinogens, adult
- Inter-human susceptibility to carcinogens, early life (including prenatal)
- Inter-human susceptibility to non-carcinogens, early life (including prenatal)
- Animal findings are relevant to humans
- Findings from one route of exposure are considered representative unless data show otherwise

---

[45] Higgins, J.P.T., and S. Green, eds. 2008. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

[46] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017. Pg. 78

[47] Id. Pp. 158-9

[48] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch 4-6

ED_004886_00001477-00016

EPA has relied on standard default values ("uncertainty" or "safety" factors) that have been applied across the board to various chemicals and health outcomes. But newer science demonstrates that EPA's typical safety factor of 10 is insufficient to account for variability due to life stage, genetics, underlying disease status, and external stressors that may be due to poverty or other difficult life conditions.

For cancer, the NAS recommended that EPA include a factor to account for human variability in response to carcinogens, as EPA's current approach inaccurately assumes that there is no variability in response. They found that a factor of 25- to 50- may account for the variability between the median individual and those with more extreme responses, and recommended 25 as a reasonable default value.[49]

Similarly, EPA should increase or add factors that address cancer and non-cancer susceptibility during early life stages. While EPA does account for increased susceptibility to genotoxicants, it does not include the prenatal period or chemicals that can influence cancer through other mechanisms. California EPA's guidance incorporates factors to account for increased susceptibility for exposures that occur prenatally for carcinogens, non-mutagenic carcinogenic agents and non-carcinogens. Their literature review on differential susceptibility to carcinogens and non-carcinogens based on age and life stage derived age adjustment values for carcinogens which include the prenatal period[50] and increased the default intraspecies uncertainty factors for non-carcinogens to 30 and 100 for specific endpoints such as asthma or neurotoxicity.[51] At a minimum, EPA should use Cal EPA's age adjustment values and intraspecies uncertainty factors for incorporating age/early life susceptibility.

In general, developmental life stages, including the fetus, infancy, and childhood, are more vulnerable to chemical exposure and toxicity. However, typical EPA age-dependent adjustment factors account for other life stages but NOT fetal exposures. Recent studies have demonstrated differential expression and activity of metabolic enzymes such as Cytochrome P450 in fetal versus adult tissue, indicating potential lifestage-dependent variability in metabolic capabilities and greater vulnerability during fetal development not accounted for in current risk assessment practices.[52] This is a critical point to address, as disruptions during fetal development have implications for health and disease in adulthood. EPA should evaluate this rich body of literature to identify the most up-to-date scientific knowledge regarding human variability and susceptibility and incorporate these scientifically-based default values in their assessments unless there are chemical-specific data supporting departing from the defaults. California EPA also developed child-specific risk values for chemicals (e.g., atrazine, lead, nickel, manganese, heptachlor) that specifically address routes of exposure and differences in susceptibility

---

[49] Id. Pg. 168

[50] California EPA 2009. Cal EPA 2009. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document for Cancer Potency Factors: Methodologies for derivation, listing of available values, and adjustments to allow for early life stage exposures. http://oehha.ca.gov/media/downloads/crnr/tsdcancerpotency.pdf

[51] Cal EPA 2008. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document For the Derivation of Noncancer Reference Exposure Levels http://oehha.ca.gov/media/downloads/crnr/noncancertsdfinal.pdf

[52] Sadler, N.C., Nandhikonda, P., Webb-Robertson, B.J., Ansong, C., Anderson, L.N., Smith, J.N., Corley, R.A. and Wright, A.T., 2016. Hepatic cytochrome P450 activity, abundance, and expression throughout human development. *Drug Metabolism and Disposition*, *44*(7), pp.984-991.

ED_004886_00001477-00017

unique to children compared to adults.[53] EPA should review these evaluations and incorporate these values as appropriate. Furthermore, a default guidance principle should be that animal findings are relevant to humans unless there is sufficient and compelling information to support otherwise.

Risk estimates

EPA should not use MOE (margin of exposure) as an analysis method in the risk evaluation process moving forward. MOE is not an estimate of risk—it is a single number that is a version of the "bright line" approach like the Reference Dose (or Reference Concentration for inhalation doses). MOE is calculated by dividing the point of departure (e.g., LOAELs, NOAELs or BMDLs) by estimated exposure values, and this 'bright line' approach does not provide information about the magnitude of the risks above, at, or below this line. Further, it implies that there is a "safe" level of exposure below which no harm will occur. While this may be true for a select few chemicals, the NAS Science and Decisions report recognizes that this is not a valid assumption for all chemicals and has recommended moving away from such "bright line" approaches which do not establish risk estimates across the full range of exposures.[54] Additionally, the MOE will not provide the necessary information for future analysis of risks and benefits that will be critical for decision-making on these chemicals.[55] We recommend that EPA utilize available analytical methods such as PODs based on a BMD to develop quantified estimates of risk.

EPA appropriately states that a dose-response assessment will be conducted for all identified human health hazard endpoints. PODs should also be developed for every endpoint unless the data are insufficient to develop a model. For calculating cancer or non-cancer risks, we recommend always using a point of departure (POD) of a benchmark dose (BMD) at 1%. The POD should be based on a BMD calculation, not the NOAEL/LOAEL, unless the data are insufficient to model. EPA already recognizes the features that make BMDs superior: BMDs account for the shape of the dose–response function; are independent of study design, such as the space between dosing; and are comparable across chemicals.[56]

Historically, for carcinogens that are direct mutagens or are associated with large human body burdens, EPA has assumed there is no threshold of effect. But the NAS Science and Decisions report highlights the science indicating that this linear presumption with no threshold is appropriate for the calculation of both cancer and non-cancer risks, and regardless of whether a carcinogen is a mutagen. For example, dose-response relationships can be linear at low dose when exposures contribute to an existing disease process, add to background processes and/ or exposures, and interact with interindividual variability or susceptibility.[57] Science and Decisions recommends harmonizing cancer and non-cancer risk assessment approaches. Therefore, for calculating non-mutagen cancer or non-cancer risks based on a POD, EPA should use the same approach as for mutagens, which assumes a straight line from the POD. In fact, a

---

[53] California Environmental Protection Agency. Office of Environmental Health Hazard Assessment (OEHHA). Child-Specific Reference Doses (chRDs) Finalized to Date. Available from: http://oehha.ca.gov/risk-assessment/chrd/table-all-chrds

[54] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

[55] McGartland A, Revesz R, Axelrad DA, Dockins C, Sutton P, Woodruff TJ. Estimating the health benefits of environmental regulations. Science (80-). 2017 Aug 4;357(6350):457–8.

[56] Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, Rusyn I. 2014. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. Environmental health perspectives. 122(5).

[57] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch.5

linear relationship may actually underestimate risks for some chemicals where the dose-response curve is supra-linear.

6. Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.

Production volumes for both asbestos and pigment violet 29 have been claimed as CBI. Production volume is basic information about a chemical to which the public and scientists should have access. We urge EPA to move forward with substantiating such claims under the new TSCA.

Appendix

When and how to update systematic reviews: consensus and checklist
Garner et al.
BMJ 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

ED_004886_00001477-00020

OPEN ACCESS

When and how to update systematic reviews: consensus and checklist

Paul Garner,[1] Sally Hopewell,[2] Jackie Chandler,[3] Harriet MacLehose,[3] Elie A Akl,[5,6] Joseph Beyene,[7] Stephanie Chang,[8] Rachel Churchill,[9] Karin Dearness,[10] Gordon Guyatt,[4] Carol Lefebvre,[11] Beth Liles,[12] Rachel Marshall,[3] Laura Martínez García,[13] Chris Mavergames,[14] Mona Nasser,[15] Amir Qaseem,[16,17] Margaret Sampson,[18] Karla Soares-Weiser,[3] Yemisi Takwoingi,[19] Lehana Thabane,[4,20] Marialena Trivella,[21] Peter Tugwell,[22] Emma Welsh,[23] Ed C Wilson,[2,4] Holger J Schünemann[4,5]

For numbered affiliations see end of article.

Correspondence to: P Garner Paul.Garner@lstmed.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

Accepted: 26 May 2016

Updating of systematic reviews is generally more efficient than starting all over again when new evidence emerges, but to date there has been no clear guidance on how to do this. This guidance helps authors of systematic reviews, commissioners, and editors decide when to update a systematic review, and then how to go about updating the review.

Systematic reviews synthesise relevant research around a particular question. Preparing a systematic review is time and resource consuming, and provides a snapshot of knowledge at the time of incorporation of data from studies identified during the latest search. Newly identified studies can change the conclusion of a review. If they have not been included, this threatens the validity of the review, and, at worst, means the review could mislead. For patients and other healthcare consumers, this means that care and policy development might not be fully informed by the latest research; furthermore, researchers could be misled and carry out research in areas where no further research is actually needed.[1] Thus, there are clear benefits to updating reviews, rather than duplicating the entire process as new evidence emerges or new methods develop. Indeed, there is probably added value to updating a review, because this will include taking into account comments and criticisms, and adoption of new methods in an iterative process.[2-6]

Cochrane has over 20 years of experience with preparing and updating systematic reviews, with the publication of over 6000 systematic reviews. However, Cochrane's principle of keeping all reviews up to date has not been possible, and the organisation has had to adapt: from updating when new evidence becomes available,[7] to updating every two years,[8] to updating based on need and priority.[9] This experience has shown that it is not possible, sensible, or feasible

SUMMARY POINTS

Updating systematic reviews is, in general, more efficient than starting afresh when new evidence emerges. The panel for updating guidance for systematic reviews (PUGs; comprising review authors, editors, statisticians, information specialists, related methodologists, and guideline developers) met to develop guidance for people considering updating systematic reviews. The panel proposed the following:

1. Decisions about whether and when to update a systematic review are judgments made for individual reviews at a particular time. These decisions can be made by agencies responsible for systematic review portfolios, journal editors with systematic review update services, or author teams considering embarking on an update of a review.
2. The decision needs to take into account whether the review addresses a current question, uses valid methods, and is well conducted; and whether there are new relevant methods, new studies, or new information on existing included studies. Given this information, the agency, editors, or authors need to judge whether the update will influence the review findings or credibility sufficiently to justify the effort in updating it.
3. Review authors and commissioners can use a decision framework and checklist to navigate and report these decisions with "update status" and rationale for this status. The panel noted that the incorporation of new synthesis methods (such as Grading of Recommendations Assessment, Development and Evaluation (GRADE)) is also often likely to improve the quality of the analysis and the clarity of the findings.
4. Given a decision to update, the process needs to start with an appraisal and revision of the background, question, inclusion criteria, and methods of the existing review.
5. Search strategies should be refined, taking into account changes in the question or inclusion criteria. An analysis of yield from the previous edition, in relation to databases searched, terms, and languages can make searches more specific and efficient.
6. In many instances, an update represents a new edition of the review, and authorship of the new version needs to follow criteria of the International Committee of Medical Journal Editors (ICMJE). New approaches to publishing licences could help new authors build on and re-use the previous edition while giving appropriate credit to the previous authors.

The panel also reflected on this guidance in the context of emerging technological advances in software, information retrieval, and electronic linkage and mining. With good synthesis and technology partnerships, these advances could revolutionise the efficiency of updating in the coming years.

ED_004886_00001477-00021

to continually update all reviews all the time. Other groups, including guideline developers and journal editors, adopt updating principles (as applied, for example, by the *Systematic Reviews* journal; https://systematicreviewsjournal.biomedcentral.com/).

The panel for updating guidance for systematic reviews (PUGs) group met to draw together experiences and identify a common approach. The PUGs guidance can help individuals or academic teams working outside of a commissioning agency or Cochrane, who are considering writing a systematic review for a journal or to prepare for a research project. The guidance could also help these groups decide whether their effort is worthwhile.

Panel selection and procedures

An international panel of authors, editors, clinicians, statisticians, information specialists, other methodologists, and guideline developers was invited to a two day workshop at McMaster University, Hamilton, Canada, on 26-27 June 2014, organised by Cochrane. The organising committee selected the panel (web appendix 1). The organising committee invited participants, put forward the agenda, collected background materials and literature, and drafted the structure of the report.

The purpose of the workshop was to develop a common approach to updating systematic reviews, drawing on existing strategies, research, and experience of people working in this area. The selection of participants aimed on broad representation of different groups involved in producing systematic reviews (including authors, editors, statisticians, information specialists, and other methodologists), and those using the reviews (guideline developers and clinicians). Participants within these groups were selected on their expertise and experience in updating, in previous work developing methods to assess reviews, and because some were recognised for developing approaches within organisations to manage updating strategically. We sought to identify general approaches in this area, and not be specific to Cochrane; although inevitably most of the panel were somehow engaged in Cochrane.

The workshop structure followed a series of short presentations addressing key questions on whether, when, and how to update systematic reviews. The proceedings included the management of authorship and editorial decisions, and innovative and technological approaches. A series of small group discussions followed each question, deliberating content, and forming recommendations, as well as recognising uncertainties. Large group, round table discussions deliberated further these small group developments. Recommendations were presented to an invited forum of individuals with varying levels of expertise in systematic reviews from McMaster University (of over 40 people), widely known for its contributions to the field of research evidence synthesis. Their comments helped inform the emerging guidance.

The organising committee became the writing committee after the meeting. They developed the guidance arising from the meeting, developed the checklist and diagrams, added examples, and finalised the manuscript. The guidance was circulated to the larger group three times, with the PUGs panel providing extensive feedback. This feedback was all considered and carefully addressed by the writing committee. The writing committee provided the panel with the option of expressing any additional comments from the general or specific guidance in the report, and the option for registering their own view that might differ to the guidance formed and their view would be recorded in an annex. In the event, consensus was reached, and the annex was not required.

Definition of update

The PUGs panel defined an update of a systematic review as a new edition of a published systematic review with changes that can include new data, new methods, or new analyses to the previous edition. This expands on a previous definition of a systematic review update.[10] An update asks a similar question with regard to the participants, intervention, comparisons, and outcomes (PICO) and has similar objectives; thus it has similar inclusion criteria. These inclusion criteria can be modified in the light of developments within the topic area with new interventions, new standards, and new approaches. Updates will include a new search for potentially relevant studies and incorporate any eligible studies or data; and adjust the findings and conclusions as appropriate. Box 1 provides some examples.

Which systematic reviews should be updated and when?

Any group maintaining a portfolio of systematic reviews as part of their normative work, such as guidelines panels or Cochrane review groups, will need to prioritise which reviews to update. Box 2 presents the approaches used by the Agency for HealthCare Research and Quality (AHRQ) and Cochrane to prioritise which systematic reviews to update and when. Clearly, the responsibility for deciding which systematic reviews should be updated and when they will be updated will vary: it may be centrally organised and resourced, as with the AHRQ scientific resource centre (box 2). In Cochrane, the decision making process is decentralised to the Cochrane Review Group editorial team, with different approaches applied, often informally.

Box 1: Examples of what factors might change in an updated systematic review

- A systematic review of steroid treatment in tuberculosis meningitis used GRADE methods and split the composite outcome in the original review of death plus disability into its two components. This improved the clarity of the reviews findings in relation to the effects and the importance of the effects of steroids on death and on disability.[11]
- A systematic review of dihydroartemisinin-piperaquine (DHAP) for treating malaria was updated with much more detailed analysis of the adverse effect data from the existing trials as a result of questions raised by the European Medicines Agency. Because the original review included other comparisons, the update required extracting only the DHAP comparisons from the original review, and a modification of the title and the PICO.[12]
- A systematic review of atorvastatin was updated with simple uncontrolled studies.[13] This update allowed comparisons with trials and strengthened the review findings.[14]

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001477-00022

The PUGs panel recommended an individualised approach to updating, which used the procedures summarised in figure 1. The figure provides a status category, and some options for classifying reviews into each of these categories, and builds on a previous decision tool and earlier work developing an updating classification system.[15 16] We provide a narrative for each step.

Step 1: assess currency

Does the published review still address a current question?

An update is only worthwhile if the question is topical for decision making for practice, policy, or research priorities (fig 1). For agencies, people responsible for managing a portfolio of systematic reviews, there is a need to use both formal and informal horizon scanning. This type of scanning helps identify questions with currency, and can help identify those reviews that should be updated. The process could include monitoring policy debates around the review, media outlets, scientific (and professional) publications, and linking with guideline developers.

Has the review had good access or use?

Metrics for citations, article access and downloads, and sharing via social or traditional media can be used as proxy or indicators for currency and relevance of the review. Reviews that are widely cited and used could be important to update should the need arise. Comparable reviews that are never cited or rarely downloaded, for example, could indicate that they are not addressing a question that is valued, and might not be worth updating.

In most cases, updated reviews are most useful to stakeholders when there is new information or methods that result in a change in findings. However, there are some circumstances in which an up to date search for information is important for retaining the credibility of the review, regardless of whether the main findings would change or not. For example, key stakeholders would dismiss a review if a study is carried out in a relevant geographical setting but is not included; if a large, high profile study that might not change the findings is not included; or if an up to date search is required for a guideline to achieve credibility. Box 3 provides such examples. If the review does not answer a current question, the intervention has been superseded, then a decision can be made not to update and no further intelligence gathering is required (fig 1).

Did the review use valid methods and was it well conducted?

If the question is current and clearly defined, the systematic review needs to have used valid methods and be well conducted. If the review has vague inclusion criteria, poorly articulated outcomes, or inappropriate methods, then updating should not proceed. If the question is current, and the review has been cited or used, then it might be appropriate to simply start with a new protocol. The appraisal should take into account the methods in use when the review was done.

Step 2: identify relevant new methods, studies, and other information

Are there any new relevant methods?

If the question is current, but the review was done some years ago, the quality of the review might not meet current day standards. Methods have advanced quickly, and data extraction and understanding of the review process have become more sophisticated. For example:

- Methods for assessing risk of bias of randomised trials,[23] diagnostic test accuracy (QUADAS-2),[24] and observational studies (ROBINS-1).[25]
- Application of summary of findings, evidence profiles, and related GRADE methods has meant the characteristics of the intervention, characteristics of the participants, and risk of bias are more thoroughly and systematically documented.[26 27]
- Integration of other study designs containing evidence, such economic evaluation and qualitative research.[28]

There are other incremental improvements in a wide range of statistical and methodological areas, for example, in describing and taking into account cluster

Box 2: Examples of how different organisations decide on updating systematic reviews

Agency for Healthcare Research and Quality (US)

The AHRQ uses a needs based approach; updating systematic reviews depends on an assessment of several criteria:

1. Stakeholder impact
   - Interest from stakeholder partners (such as consumers, funders, guideline developers, clinical societies, James Lind Alliance)
   - Use and uptake (for example, frequency of citations and downloads)
   - Citation in scientific literature including clinical practice guidelines
2. Currency and need for update
   - New research is available
   - Review conclusions are probably dated
3. Update decision
   - Based on the above criteria, the decision is made to either update, archive, or continue surveillance.

Cochrane

Of over 50 Cochrane editorial teams, most but not all have some systems for updating, although this process can be informal and loosely applied. Most editorial teams draw on some or all of the following criteria:

1. Strategic importance
   - Is the topic a priority area (for example, in current debates or considered by guidelines groups)?
   - Is there important new information available?
2. Practicalities in organising the update that many groups take into account
   - Size of the task (size and quality of the review, and how many new studies or analyses are needed)
   - Availability and willingness of the author team
3. Impact of update
   - New research impact on findings and credibility
   - Consider whether new methods will improve review quality
4. Update decision
   - Priority to update, postpone update, class review as no longer requiring an update

ED_004886_00001477-00023

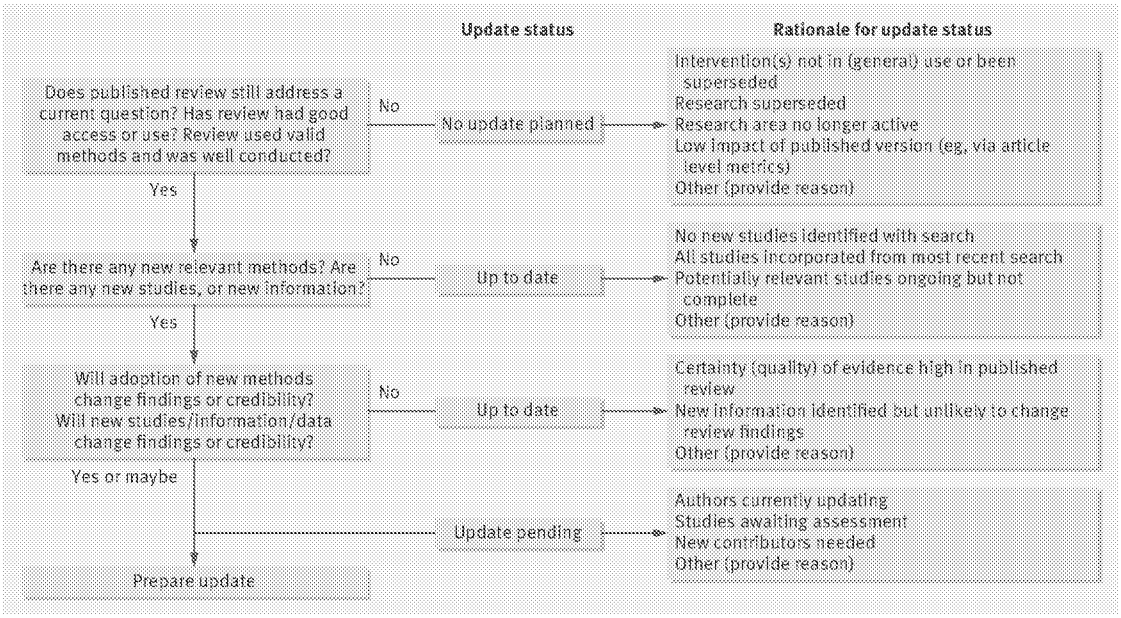


Fig 1 | Decision framework to assess systematic reviews for updating, with standard terms to report such decisions

randomised trials.[29] AMSTAR can assess the overall quality of a systematic review,[30] and the ROBIS tool can provide a more detailed assessment of the potential for bias.[31]

Are there any new studies or other information?

If an authoring or commissioning team wants to ensure that a particular review is up to date, there is a need for routine surveillance for new studies that are potentially relevant to the review, by searching and trial register inspection at regular intervals. This process has several approaches, including:

- Formal surveillance searching[32]
- Updating the full search strategies in the original review and running the searches
- Tracking studies in clinical trial and other registers
- Using literature appraisal services[33]
- Using a defined abbreviated search strategy for the update[34]
- Checking studies included in related systematic reviews.[35]

How often this surveillance is done, and which approaches to use, depend on the circumstances and the topic. Some topics move quickly, and the definition of "regular intervals" will vary according to the field and according to the state of evidence in the field. For example, early in the life of a new intervention, there might be a plethora of studies, and surveillance would be needed more frequently.

Box 3: Examples of a systematic review's currency

- The public is interested in vitamin C for preventing the common cold: the Cochrane review includes over 29 trials with either no or small effects, concluding good evidence of no important effects.[17] Assessment: still a current question for the public.
- Low osmolarity oral rehydration salt (ORS) solution versus standard solution for acute diarrhoea in children: the 2001 Cochrane review[18] led the World Health Organization to recommend ORS solution formula worldwide to follow the new ORS solution formula[19] and this has now been accepted globally. Assessment: no longer a current question.
- Routine prophylactic antibiotics with caesarean section: the Cochrane review reports clear evidence of maternal benefit from placebo controlled trials but no information on the effects on the baby.[20] Assessment: this is a current question.
- A systematic review published in the *Lancet* examined the effects of artemisinin based combination treatments compared with monotherapy for treating malaria and showed clear benefit.[21] Assessment: this established the treatment globally and is no longer a current question and no update is required.
- A Cochrane review of amalgam restorations for dental caries[22] is unlikely to be updated because the use of dental amalgam is declining, and the question is not seen as being important by many dental specialists. Assessment: no longer a current question.

Step 3: assess the effect of updating the review

Will the adoption of new methods change the findings or credibility?

Editors, referees, or experts in the topic area or methodologists can provide an informed view of whether a review can be substantially improved by application of current methodological expectations and new methods (fig 1). For example, a Cochrane review of iron supplementation in malaria concluded that there was "no significant difference between iron and placebo detected."[36] An update of the review included a GRADE assessment of the certainty of the evidence, and was able to conclude with a high degree of certainty that iron does not cause an excess of clinical malaria because the upper relative risk confidence intervals of harm was 1.0 with high certainty of evidence.[37]

Will the new studies, information, or data change the findings or credibility?

The assessment of new data contained in new studies and how these data might change the review is often used to determine whether an update should go ahead,

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001477-00024

Table 1 | Formal prediction tools: how potentially relevant new studies can affect review conclusions

Method	Description of approach	How it could be used	Advantages	Limitations	Validation
GRADE approach[38]	Considers whether the evidence certainty might change in the update (for example, because of lack of high certainty evidence, or because new evidence contradicts existing high certainty evidence). High certainty of evidence for critical outcomes could lower the priority for updating. Uncertainty in the review findings increases the need to include new studies[38]	Provides a benchmark by outcome to assess whether a new trial will improve the certainty of the evidence	Pragmatic. Many reviews already include GRADE	Requires GRADE to have been used in existing review or to complete an assessment according to GRADE	GRADE summary of findings tables or evidence profiles widely validated. Use of GRADE approach to prioritising updates requires further validation
Ottawa method[39-41]	A simple PubMed search (using the three largest and three most recent trials from the original review) to identify new studies. If new studies are found, then the method uses quantitative signals (eg, change in significance, effect size) to assess the likelihood that the new studies will change the review conclusion, thus triggering an update	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions; reasonable agreement with RAND method[39 42-44]
RAND method[45]	An abbreviated search of five major journals to identify new studies, and a search of the US Food and Drug Administration website and external expert judgment to determine the currency of the report findings	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions, and compares well with the Ottawa method[39 43 44]
Statistical prediction tool[15]	A multicomponent decision tool to assess whether there might be any new studies for the update. If new studies are identified, a statistical prediction tool estimates the probability that this will change the review conclusion	Ranks multiple systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires commercial software	Approach validated internally[46]; requires further external validation
Value of information analysis[47 48]	Builds on the statistical prediction tool approach[15] comparing the expected health gain from new evidence with its expected cost. The gain is calculated in terms of a reduction in expected loss from reduced uncertainty and the cost is measured in days required to update the review.	Ranks selected systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires some statistical knowledge	Approach validated internally; requires further external validation

Box 4: Examples of new information other than new trials being important

- The iconic Cochrane review of steroids in preterm labour was thought to provide evidence of benefit in infants, and this question no longer required new trials. However, a new large trial published in the *Lancet* in 2015 showed that in low and middle income countries, strategies to promote the uptake of neonatal steroids increased neonatal mortality and suspected maternal infection.[49] This information needs to somehow be incorporated into the review to maintain its credibility.
- A Cochrane review of community deworming in developing countries indicates that in recent studies, there is little or no effect.[50] The inclusion of a large trial of two million children confirmed that there was no effect on mortality. Although the incorporation of the trial in the review did not change the review's conclusions, the trial's absence would have affected the credibility of the review, so it was therefore updated.
- A new paper reporting long term follow-up data on anthracycline chemotherapy as part of cancer treatment was published. Although the effects from the outcomes remained essentially unchanged, apart from this longer follow-up, the paper also included information about the performance bias in the original trial, shifting the risk of bias for several outcomes from "unknown" to "high" in the Cochrane review.[51]

and the speed with which the update should be conducted. The appraisal of these new data can be carried out in different ways. Initially, methods focused on statistical approaches to predict an overturning of the current review findings in terms of the primary or desired outcome (table 1). Although this aspect is important, additional studies can add important information to a review, which is more than just changing the primary outcome to a more accurate and reliable estimate. Box 4 gives examples.

Reviews with a high level of certainty in the results (that is, when the GRADE assessment for the body of evidence is high) are less likely to change even with the addition of new studies, information, or data, by definition. GRADE can help guide priorities in whether to update, but it is still important to assess new studies that might meet the inclusion criteria. New studies can show unexpected effects (eg, attenuation of efficacy) or provide new information about the effects seen in different circumstances (eg, groups of patients or locations).

Other tools are specifically designed to help decision making in updating. For example, the Ottawa[39] and RAND[45] methods focus on identification of new evidence, the statistical predication tool[15] calculates the probability of new evidence changing the review conclusion, and the value of information analysis approach[52] calculates the expected health gain (table 1). As yet, there has been limited external validation of these tools to determine which approach would be most effective and when.

If potentially relevant studies are identified that have not previously been assessed for inclusion, authors or those managing the updating process need to assess whether including them might affect the conclusions of the review. They need to examine the weight and certainty of the new evidence to help determine whether an update is needed and how urgent that update is. The updating team can assess this informally by judging whether new studies or data are likely to substantively affect the review, for example, by altering the certainty in an existing comparison, or by generating new comparisons and analyses in the existing review.

New information can also include fresh follow-up data on existing included studies, or information on how the studies were carried out. These should be assessed in terms of whether they might change the review findings or improve its credibility (fig 1). Indeed, if any study has been retracted, it is important the authors assess the reasons for its retraction. In the case of data fabrication, the study needs to be removed

ED_004886_00001477-00025

from the analysis and this recorded. A decision needs to be made as to whether other studies by the same author should be removed from the review and other related reviews. An investigation should also be initiated following guidelines from the Committee on Publication Ethics (COPE). Additional published and unpublished data can become available from a wide range of sources—including study investigators, regulatory agencies and industry—and are important to consider.

Preparing for an update

Refresh background, objectives, inclusion criteria, and methods

Before including new studies in the review, authors need to revisit the background, objectives, inclusion criteria, and methods of the current review. In Cochrane, this is referred to as the protocol, and editors are part of this process. The update could range from simply endorsing the current question and inclusion criteria, through to full rewriting of the question, inclusion criteria and methods, and republishing the protocol. As a field progresses with larger and better quality trials rigorously testing the questions posed, it may be appropriate to exclude weaker study designs (such as quasi-randomised comparisons or very small trials) from the update (table 2). The PUGs panel recommended that a protocol refresh will require the authors to use the latest accepted methods of synthesis, even if this means repeating data extraction for all studies.

New authors and authorship

Updated systematic reviews are new publications with new citations. An authorship team publishing an update in a scientific or medical journal is likely to manage the new edition of a review in the same way as with any other publication, and follow the ICMJE authorship criteria.[56] If the previous author or author team steps down, then they should be acknowledged in the new version. However, some might perceive that their efforts in the first version warrant continued authorship, which may be valid. The management of authorship between versions can sometimes be complicated. At worst, it delays new authors completing an update and leads to long authorship lists of people from previous versions who probably do not meet ICMJE authorship criteria. One approach with updates including new authors is to have an opt-in policy for the existing authors: they can opt in to the new edition, provided that they make clear their contribution, and this is then agreed with the entire author team.

Although they are new publications, updates will generally include content from the published version. Changing licensing rights around systematic reviews to allow new authors of future updates to remix, tweak, or build on the contributions of the original authors of the published version (similar to the rights available via a Creative Commons licence; https://creativecommons.org) could be a more sustainable and simpler approach. This approach would allow systematic reviews to continue to evolve and build on the work of a range of authors over time, and for contributors to be given credit for contributions to this previous work.

Efficient searching

In performing an update, a search based on the search conducted for the original review is required. The updated search strategy will need to take into account changes in the review question or inclusion criteria, for example, and might be further adjusted based on knowledge of running the original search strategy. The search strategy for an update need not replicate the original search strategy, but could be refined, for example, based on an analysis of the yield of the original search. These new search approaches are currently undergoing formal empirical evaluation, but they may well provide much more efficient search strategies in the future. Some examples of these possible new methods for review updates are described in web appendix 2.

In reporting the search process for the update, investigators must ensure transparency for any previous versions and the current update, and use an adapted flow diagram based on PRISMA reporting (preferred reporting items for systematic reviews and meta-analyses).[57] The search processes and strategies for the update must be adequately reported such that they could be replicated.

Table 2 | Refresh background, objectives, inclusion criteria, and methods

Protocol section	Appraisal points
Background and research question	• Review and update background section, including supporting references to take account of any changes that may have occurred. This should include updating any new information and current policy debates on the topic. • Assess whether the current review question remains relevant to patients and practice.
Inclusion criteria	• Consider whether the existing PICO(s) remain(s) current, in the light of new knowledge. • Identify any new understanding of definition of patient populations. • Identify new interventions, or those that have been withdrawn, that are no longer in use. • Identify any changes in usual care standards. • Check for standardised core outcomes sets, such as those developed in collaboration with the core outcome measures in effectiveness trials (COMET) initiative (www.comet-initiative.org) or by guideline groups since the original review. • Check for any relevant patient reported outcomes to include subsequent to the original review. • Consider any new studies with less risk of bias that might warrant a stricter study design inclusion criteria (where the older version, when there was a dearth of evidence, included observational or quasi-randomised comparisons).
Methods	• Appraise and update the methods pending relevant methodological advancements or developments. For example, if (1) there are new tools for assessing the risk of bias of individual studies or appraising the quality of a body of evidence (eg, GRADE); or (2) new and efficient search approaches are feasible, such as a targeted approach to searching, taking into account the quality of the original search, and ensuring that the search for the update is of high quality. • Update or include a summary of findings table, which is recommended for all systematic reviews, because it improves the clarity, understanding, and interpretation of the findings of a systematic review, and rapidly reduces the amount of time readers require to find key information.[53-55]

ED_004886_00001477-00026

Peer review

Systematic reviews published for the first time in peer reviewed journals are by definition peer reviewed, but practice for updates remains variable, because an update might have few changes (such as an updated search but no new studies found and therefore included) or many changes (such as revise methods and inclusion of several new studies leading to revised conclusions). Therefore, and to use peer reviewers' time most effectively, editors need to consider when to peer review an update and the type of peer reviewer most useful for a particular update (for example, topic specialist, methodologist). The decision to use peer review, and the number and expertise of the peer reviewers could depend on the nature of the update and the extent of any changes to the systematic review as part of an editor assessment. A change in the date of the search only (where no new studies were identified) would not require peer review (except, arguably, peer review of the search), but the addition of studies that lead to a change in conclusions or significant changes to the methods would require peer review. The nature of the peer review could be described within the published article.

Reporting changes

Authors should provide a clear description of the changes in approach or methods between different editions of a review. Also, authors need to report the differences in findings between the original and updated edition to help users decide how to use the new edition. The approach or format used to present the differences in findings might vary with the target user group.[58] Publishers need to ensure that all previous versions of the review remain publically accessible.

Updates can range from small adjustments to reviews being completely rewritten, and the PUGs panel spent some time debating whether the term "new edition" would be a better description than "update." However, the word "update" is now in common parlance and changing the term, the panel judged, could cause confusion. However, the debate does illustrate that an update could represent a review that asks a similar question but has been completely revised.

Technology and innovation

The updating of systematic review is generally done manually and is time consuming. There are opportunities to make better use of technology to streamline the updating process and improve efficiency (table 3). Some of these tools already exist and are in development or in early use, and some are commercially available or freely available. The AHRQ's evidence based practice centre team has recently published tools for searching and screening, and will provide an assessment of the use, reliability, and availability of these tools.[63]

Other developments, such as targeted updates that are performed rapidly and focus on updating only key components of a review, could provide different approaches to updating in the future and are being piloted and evaluated.[64] With implementation of these various innovations, the longer term goal is for "living" systematic reviews, which identify and incorporate information rapidly as it evolves over time.[60]

Concluding remarks

Updating systematic reviews, rather than addressing the same question with a fresh protocol, is generally

Table 3 | Technological innovations to improve the efficiency of updating systematic reviews

Innovation	Description	Application	Examples of software and projects,* and current status
Integrated software	Integration of applying inclusion criteria, review management systems, statistical packages, and GRADE	To facilitate greater efficiencies in review production, including their updates	Covidence (www.covidence.org): free/pay† EPPI reviewer (http://eppi.ioe.ac.uk/cms/Default.aspx): pay DistillerSR (https://distillercer.com/products/distillersr-systematic-review-software/): pay Cochrane Review Manager (http://tech.cochrane.org/revman): free/pay† GRADEpro GDT (http://gradepro.org/): free Rayyan (http://rayyan.qcri.org/): free
Systematic review data repositories	Repositories store information from review (eg, data abstraction forms and the evidence tables)	Improve updating efficiency for new or existing teams as the data abstraction forms, evidence tables, and populating data from the original review are available	Agency of Health Care Research and Quality systematic review data repository (http://srdr.ahrq.gov): operational GRADE database of evidence profiles and evidence to decision frameworks (http://dbep.gradepro.org/): operational
Semi-automation	Machine learning techniques to use alongside human efforts	Finding studies and extracting data could benefit from semi-automation creating time efficiencies[59]	RobotReviewer (http://vortext.systems/robotreviewer): free Cochrane project transform—identification of studies (http://community.cochrane.org/tools/project-coordination-and-support/transform)
Crowdsourcing	Use of volunteers to assist systematic review authors with discrete tasks	Individuals from the "crowd" assist with tasks (identifying and screening studies, translating articles, data extraction) to help in new review production and updates[60-62]	Cochrane Project Transform—crowdsourcing (link as above)
Publication linkage	Ability to link trial registration, trial publications, and reviews citing them will help transparency	This initiative could help identify studies for systematic reviews and could also show the relation between systematic review updates	A cross publisher initiative, CrossRef, is coordinating a threaded publications/linked clinical trial reports initiative to link a clinical trial report (with a trial registration number) report and derivative publications, including reviews (www.crossref.org): operational
Data linkage	Increase links between data, existing software, and reviews	To improve identification and reuse of data for review production and dissemination	Cochrane (http://linkeddata.cochrane.org/): proof of principle example at production phase, but mostly linkage projects at exploratory phase

*Further information can be located on the SR Toolbox site (http://systematicreviewtools.com/).
†Free to Cochrane contributors; other users pay.

ED_004886_00001477-00027

more efficient and allows incremental improvement over time. Mechanical rules appear unworkable, but there is no clear unified approach on when to update, and how implement this. This PUGs panel of authors, editors, statisticians, information specialists, other methodologists, and guideline developers brought together current thinking and experience in this area to provide guidance.

Decisions about whether and when to update a systematic review are judgments made at a point in time. They depend on the currency of the question asked, the need for updating to maintain credibility, the availability of new evidence, and whether new research or new methods will affect the findings.

Whether the review uses current methodological standards is important in deciding if the update will influence the review findings, quality, reliability, or credibility sufficiently to justify the effort in updating it. Those updating systematic reviews to author clinical practice guidelines might consider the influence of new study results in potentially overturning the conclusions of an existing review. Yet, even in cases where new study findings do not change the primary outcome measure, new studies can carry important information about subgroup effects, duration of treatment effects, and other relevant clinical information, enhancing the currency and breadth of review results.

An update requires appraisal and revision of the background, question, inclusion criteria, and methods of the existing review and the existing certainty in the evidence. In particular, methods might need to be updated, and search strategies reconsidered. Authors of updates need to consider inputs to the current edition, and follow ICMJE criteria regarding authorship.[56]

The PUGs panel proposed a decision framework (fig 1), with terms and categories for reporting the decisions made for updating procedures for adoption by Cochrane and other stakeholders. This framework includes journals publishing systematic review updates and independent authors considering updates of existing published reviews. The panel developed a checklist to help judgements about when and how to update.

The current emphasis of authors, guideline developers, Cochrane, and consequently this guidance has been on effects reviews. The checklists and guidance here still applies to other types of systematic reviews, such as those on diagnostic test accuracy, and this guidance will need adapting. Accumulative experience and methods development in reviews other than those of effects are likely to help refine guidance in the future.

This guidance could help groups identify and prioritise reviews for updating and hence use their finite resources to greatest effect. Software innovation and new management systems are being developed and in early use to help streamline review updates in the coming years.

AUTHOR AFFILIATIONS

[1]Cochrane Infectious Diseases Group, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK

[2]Oxford Clinical Trials Research Unit, University of Oxford, Oxford, UK

[3]Cochrane Editorial Unit, Cochrane Central Executive, London, UK

[4]Department of Clinical Epidemiology and Biostatistics and Department of Medicine, McMaster University, Hamilton, ON, Canada

[5]Cochrane GRADEing Methods Group, Ottawa, ON, Canada

[6]Department of Internal Medicine, American University of Beirut, Beirut, Lebanon

[7]Department of Mathematics and Statistics, McMaster University

[8]Evidence-based Practice Center Program, Agency for Healthcare and Research Quality, Rockville, MD, USA

[9]Centre for Reviews and Dissemination, University of York, York, UK

[10]Cochrane Upper Gastrointestinal and Pancreatic Diseases Group, Hamilton, ON, Canada

[11]Lefebvre Associates, Oxford, UK

[12]Kaiser Permanente National Guideline Program, Portland, OR, USA

[13]Iberoamerican Cochrane Centre, Barcelona, Spain

[14]Cochrane Informatics and Knowledge Management, Cochrane Central Executive, Freiburg, Germany

[15]Plymouth University Peninsula School of Dentistry, Plymouth, UK

[16]Department of Clinical Policy, American College of Physicians,Philadelphia, PA, USA

[17]Guidelines International Network, Pitlochry, UK

[18]Children's Hospital of Eastern Ontario, Ottawa, ON, Canada

[19]Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[20]Biostatistics Unit, Centre for Evaluation, McMaster University, Hamilton, ON, Canada

[21]Centre for Statistics in Medicine, University of Oxford, Oxford, UK

[22]University of Ottawa, Ottawa, ON, Canada

[23]Cochrane Airways Group, Population Health Research Institute, St George's, University of London, London, UK

[24]Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, UK

Contributors: HJS initiated the workshop. JC, SH, PG, HM, and HJS organised the materials and the agenda. SH wrote up the proceedings. PG wrote the paper from the proceedings and coordinated the development of the final guidance; JC, SH, HM, and HJS were active in the finalising of the guidance. All PUGs authors contributed to three rounds of manuscript revision.

Funding: Attendance at this meeting, for those attendees not directly employed by Cochrane, was not funded by Cochrane beyond the reimbursement of out of pocket expenses for those attendees for whom this was appropriate. Expenses were not reimbursed for US federal government attendees, in line with US government policy. Statements in the manuscript should not be construed as endorsement by the US Agency for Healthcare Research and Quality or the US Department of Health and Human Services.

Competing interests: All participants have a direct or indirect interest in systematic reviews and updating as part of their job or academic career. Most participants contribute to Cochrane, whose mission includes a commitment to the updating of its systematic review portfolio. JC, HM, RM, CM, KS-W, and MT are, or were at that time, employed by the Cochrane Central Executive.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 3.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/3.0/.

1 Shekelle PG, Ortiz E, Rhodes S, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated?*JAMA* 2001;286:1461-7. doi:10.1001/jama.286.12.1461.

2 Claxton K, Cohen JT, Neumann PJ. When is evidence sufficient?*Health Aff (Millwood)* 2005;24:93-101. doi:10.1377/hlthaff.24.1.93.

3 Fenwick E, Claxton K, Sculpher M, et al. *Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Paper 179.*Centre for Health Economics, University of York, 2000.

4 Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty—when is there sufficient evidence?*Value Health* 2005;8:433-46. doi:10.1111/j.1524-4733.2005.00033.x.

ED_004886_00001477-00028

5 Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *J Health Serv Res Policy* 1997;2:26-30.
6 Wilson E, Abrams K. From evidence based economics to economics based evidence: using systematic review to inform the design of future research. In: Shemilt I, Mugford M, Vale L, et al. eds. *Evidence based economics*.Blackwell Publishing, 2010doi:10.1002/9781444320398.ch12.
7 Chalmers I, Enkin M, Keirse MJ. Preparing and updating systematic reviews of randomized controlled trials of health care. *Milbank Q* 1993;71:411-37. doi:10.2307/3350409.
8 Higgins J, Green S, Scholten R. Chapter 3. *Maintaining reviews: updates, amendments and feedback: Version 5.1.0 (updated March 2011)*.Cochrane Collaboration, 2011.
9 Cochrane. Editorial and publishing policy resource. http://community.cochrane.org/editorial-and-publishing-policy-resource. 2016.
10 Moher D, Tsertsvadze A. Systematic reviews: when is an update an update?*Lancet* 2006;367:881-3. doi:10.1016/S0140-6736(06)68358-X.
11 Prasad K, Singh MB, Ryan H. Corticosteroids for managing tuberculous meningitis. *Cochrane Database Syst Rev* 2016;4:CD002244.
12 Zani B, Gathu M, Donegan S, Olliaro PL, Sinclair D. Dihydroartemisinin-piperaquine for treating uncomplicated *Plasmodium falciparum* malaria. *Cochrane Database Syst Rev* 2014;1:CD010927.
13 Adams SP, Tsang M, Wright JM. Lipid lowering efficacy of atorvastatin. *Cochrane Database Syst Rev* 2012;12:CD008226.
14 Higgins J. Convincing evidence from controlled and uncontrolled studies on the lipid-lowering effect of a statin. *Cochrane Database Syst Rev* 2012;12:ED000049.
15 Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ* 2013;347:f7191. doi:10.1136/bmj.f7191.
16 MacLehose H, Hilton J, Tovey D, et al. The Cochrane Library: revolution or evolution? Shaping the future of Cochrane content. Background paper for The Cochrane Collaboration's Strategic Session Paris, France, 18 April 2012. http://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/2012-CC-strategic-session_full-report.pdf.
17 Hemilä H, Chalker E. Vitamin C for preventing and treating the common cold. *Cochrane Database Syst Rev* 2013;(1):CD000980.
18 Hahn S, Kim S, Garner P. Reduced osmolarity oral rehydration solution for treating dehydration caused by acute diarrhoea in children. *Cochrane Database Syst Rev* 2002;(1):CD002847.
19 World Health Organization (WHO). Reduced osmolarity oral rehydration salts (ORS) formulation. A report from a meeting of Experts jointly organized by UNICEF and WHO. New York: Child and Adolescent Health and Development, 18 July 2001 http://apps.who.int/iris/bitstream/10665/67322/1/WHO_FCH_CAH_01.22.pdf.
20 Smaill FM, Grivell RM. Antibiotic prophylaxis versus no prophylaxis for preventing infection after cesarean section. *Cochrane Database Syst Rev* 2014;(10):CD007482.
21 Adjuik M, Babiker A, Garner P, Olliaro P, Taylor W, White N. International Artemisinin Study Group. Artesunate combinations for treatment of malaria: meta-analysis. *Lancet* 2004;363:9-17. doi:10.1016/S0140-6736(03)15162-8.
22 Agnihotry A, Fedorowicz Z, Nasser M. Adhesively bonded versus non-bonded amalgam restorations for dental caries. *Cochrane Database Syst Rev* 2016;3:CD007517.
23 Higgins JP, Altman DG, Gøtzsche PC, et al. Cochrane Bias Methods Group Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928.
24 Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36. doi:10.7326/0003-4819-155-8-201110180-00009.
25 Sterne JAC, Higgins JPT, Reeves BC; on behalf of the development group for ROBINS-I. A tool for assessing risk of bias in non-randomized studies of interventions, version 7. March 2016. www.riskofbias.info.
26 Guyatt GH, Oxman AD, Vist GE, et al. GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD.
27 Schünemann HJ. Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision?*J Clin Epidemiol* 2016;75:6-15. doi:10.1016/j.jclinepi.2016.03.018.
28 Gough D. Qualitative and mixed methods in systematic reviews. *Syst Rev* 2015;4:181. doi:10.1186/s13643-015-0151-y.
29 Richardson M, Garner P, Donegan S. Cluster randomised trials in Cochrane reviews: evaluation of methodological and reporting practice. *PLoS One* 2016;11:e0151818. doi:10.1371/journal.pone.0151818.
30 Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10. doi:10.1186/1471-2288-7-10.
31 Whiting P, Savović J, Higgins JP, et al. ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005.
32 Sampson M, Shojania KG, McGowan J, et al. Surveillance search techniques identified the need to update systematic reviews. *J Clin Epidemiol* 2008;61:755-62. doi:10.1016/j.jclinepi.2007.10.003.
33 Hemens BJ, Haynes RB. McMaster Premium LiteratUre Service (PLUS) performed well for identifying new studies for updated Cochrane reviews. *J Clin Epidemiol* 2012;65:62-72.e1. doi:10.1016/j.jclinepi.2011.02.010.
34 Sagliocca L, De Masi S, Ferrigno L, Mele A, Traversa G. A pragmatic strategy for the review of clinical evidence. *J Eval Clin Pract* 2013;19:689-96. doi:10.1111/jep.12020.
35 Rada G, Peña J, Capurro D, et al. How to create a matrix of evidence in Epistemonikos. Abstracts of the 22nd Cochrane Colloquium; Evidence-informed public health: opportunities and challenges; Hyderabad, India. *Cochrane Database Syst Rev* 2014;suppl 1:132.
36 Okebe JU, Yahav D, Shbita R, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2011;(10):CD006589.
37 Neuberger A, Okebe J, Yahav D, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2016;2:CD006589.
38 Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401-6. doi:10.1016/j.jclinepi.2010.07.015.
39 Chung M, Newberry SJ, Ansari MT, et al. Two methods provide similar signals for the need to update systematic reviews. *J Clin Epidemiol* 2012;65:660-8. doi:10.1016/j.jclinepi.2011.12.004.
40 Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med* 2007;147:224-33. doi:10.7326/0003-4819-147-4-200708210-00179.
41 Shojania K, Sampson M, Ansari M, et al. *Updating systematic reviews; AHRQ technical reviews; report no 07-0087*.Agency for Healthcare Research and Quality, 2007.
42 Pattanittum P, Laopaiboon M, Moher D, Lumbiganon P, Ngamjarus C. A comparison of statistical methods for identifying out-of-date systematic reviews. *PLoS One* 2012;7:e48894. doi:10.1371/journal.pone.0048894.
43 Shekelle PG, Motala A, Johnsen B, Newberry SJ. Assessment of a method to detect signals for updating systematic reviews. *Syst Rev* 2014;3:13. doi:10.1186/2046-4053-3-13.
44 Shekelle PG, Newberry SJ, Wu H, et al. *Identifying signals for updating systematic reviews: a comparison of two methods; report no 11-EHC042-EF*.Agency for Healthcare Research and Quality, 2011.
45 Shekelle P, Newberry S, Maglione M, et al. *Assessment of the need to update comparative effectiveness reviews: report of an initial rapid program assessment (2005-2009)*.Agency for Healthcare Research and Quality, 2009.
46 Tovey D, Marshall R, Bazian, Hopewell S, Rader T. Fit for purpose: centralised updating support for high-priority Cochrane Reviews; National Institute for Health Research Cochrane-National Health Service Engagement Award Scheme, July 2011. https://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/10_4000_01%20Fit%20for%20purpose%20-%20centralised%20updating%20support%20for%20high%20priority%20Cochrane%20Reviews%20FINAL%20REPORT.pdf.
47 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341-64. doi:10.1016/S0167-6296(98)00039-3.
48 Wilson EC. A practical guide to value of information analysis. *Pharmacoeconomics* 2015;33:105-21. doi:10.1007/s40273-014-0219-x.
49 Althabe F, Belizán JM, McClure EM, et al. A population-based, multifaceted strategy to implement antenatal corticosteroid treatment versus standard care for the reduction of neonatal mortality due to preterm birth in low-income and middle-income countries: the ACT cluster-randomised trial. *Lancet* 2015;385:629-39. doi:10.1016/S0140-6736(14)61651-2.
50 Taylor-Robinson D, Maayan N, Soares-Weiser K, et al. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin and school performance. *Cochrane Database Syst Rev* 2012;(11):CD000371.
51 van Dalen EC, van der Pal HJ, Kremer LC. Different dosage schedules for reducing cardiotoxicity in people with cancer receiving anthracycline chemotherapy. *Cochrane Database Syst Rev* 2016;3:CD005008.
52 Wilson E. on behalf of the Cochrane Priority Setting and Campbell & Cochrane Economics Methods Groups. Which study when? Proof of concept of a proposed automated tool to help decision which reviews to update first. *Cochrane Database Syst Rev* 2014;suppl 2:29-31.
53 Rosenbaum SE, Glenton C, Nylund HK, Oxman AD. User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews. *J Clin Epidemiol* 2010;63:607-19. doi:10.1016/j.jclinepi.2009.12.013.
54 Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. *J Clin Epidemiol* 2010;63:620-6. doi:10.1016/j.jclinepi.2009.12.014.

ED_004886_00001477-00029

55 Vandvik PO, Santesso N, Akl EA, et al. Formatting modifications in GRADE evidence profiles improved guideline panelists comprehension and accessibility to information. A randomized trial. *J Clin Epidemiol* 2012;65:748-55. doi:10.1016/j.jclinepi.2011.11.013.
56 International Committee of Medical Journal Editors (ICMJE). Defining the role of authors and contributors. 2016. www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.
57 Stovold E, Beecher D, Foxlee R, Noel-Storr A. Study flow diagrams in Cochrane systematic review updates: an adapted PRISMA flow diagram. *Syst Rev* 2014;3:54. doi:10.1186/2046-4053-3-54.
58 Newberry SJ, Shekelle PG, Vaiana M, et al. *Reporting the findings of updated systematic reviews of comparative effectiveness: how do users want to view new information? report no 13-EHC093-EF* Agency for Healthcare Research and Quality, 2013.
59 Marshall IJ, Kuiper J, Wallace BC. Automating risk of bias assessment for clinical trials BCB'14. Proceedings of the 5th ACM conference on Bioinformatics, computational biology, and health informatics. 2014:88-95. http://thirdworld.nl/automating-risk-of-bias-assessment-for-clinical-trials.
60 Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014;11:e1001603. doi:10.1371/journal.pmed.1001603.
61 Elliott J, Sim I, Thomas J, et al. #CochraneTech: technology and the future of systematic reviews. *Cochrane Database Syst Rev* 2014;(9):ED000091.
62 Cochrane. Project transform: the Cochrane Collaboration. 2016. http://community.cochrane.org/tools/project-coordination-and-support/transform.
63 Paynter R, Bañez L, Berlinerm E, et al. EPC methods: an exploration of the use of text-mining software in systematic reviews. Research white paper. AHRQ publication 16-EHC023-EF. Agency for Healthcare Research and Quality, April 2016. https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2214.
64 Soares-Weiser K, Marshall R, Bergman H, et al. Updating Cochrane Reviews: results of the first pilot of a focused update. *Cochrane Database Syst Rev* 2014;suppl 1:31-3.

© BMJ Publishing Group Ltd 2016

Web appendix: Supplementary material

September 19, 2017

Comments from Academics, Scientists and Clinicians on the Risk Evaluation Scoping Efforts Under TSCA for Ten Chemical Substances

Submitted online via *Regulations.gov* to dockets: EPA-HQ-OPPT-2016-0736, EPA-HQ-OPPT-2016-0741, EPA-HQ-OPPT-2016-0733, EPA-HQ-OPPT-2016-0723, EPA-HQ-OPPT-2016-0735, EPA-HQ-OPPT-2016-0742, EPA-HQ-OPPT-2016-0743, EPA-HQ-OPPT-2016-0732, EPA-HQ-OPPT-2016-0725, EPA-HQ-OPPT-2016-0737

These comments are submitted on behalf of the undersigned academic, scientists, and clinicians. We declare collectively that we have no direct or indirect financial or fiduciary interest in any chemical under consideration in these risk evaluations. The co-signers' institutional affiliations are included for identification purposes only and do not necessarily imply any institutional endorsement or support, unless indicated otherwise.

We appreciate the opportunity to provide written comments on the scope of risk evaluations for the first ten chemical substances for risk evaluations pursuant to the Toxic Substances Control Act (TSCA), as amended by the Frank R. Lautenberg Chemical Safety of the 21st Century Act (Lautenberg TSCA). Collectively, these chemicals represent an aggregate production volume of more than 1 billion pounds a year in 2015.[1] Some of these chemicals have assessments, and in some cases even restrictions, under other federal programs – but none of these other programs has the mandate given to EPA under the new TSCA: to comprehensively evaluate chemicals and ensure that they do not pose an unreasonable risk to human health and the environment, with special consideration to those most vulnerable amongst us. Therefore, the task ahead for EPA is critical.

These first ten evaluations are also consequential because they will be precedent setting for the implementation of evaluation of science under TSCA. The consequent health impacts of EPA's decisions – for better or worse – will be borne by generations of American children, workers, families, and communities. With so much at stake, we welcome EPA's engagement with the public in this process and we offer EPA concrete approaches to embed the most current scientific principles in its methods to assess the hazards and risks of environmental chemicals.

Our comments address the following main points:

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**
2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**
3. **EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations**

[1] This is the aggregate production volume estimate for the 8 chemicals with production volume information available. For asbestos and pigment violet 29, manufacturers/ importers claimed production volumes as confidential business information (CBI).

ED_004886_00001520-00001

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**
5. **For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.**
6. **Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.**

We are appreciative of the opportunity to provide public input and we look forward to continuing to participate in such opportunities in the near future. Please do not hesitate to contact us with any questions regarding these comments.

Sincerely,

Veena Singla, PhD; Juleen Lam, PhD; and Tracey Woodruff, PhD
Program on Reproductive Health and the Environment
University of California, San Francisco

Patricia D. Koman, MPP, PhD
President and Senior Health Scientist
Green Barn Research

Kathy Attar, MPH
Toxics Program Manager
Physicians for Social Responsibility

Sheila Brear, BDS
Associate Dean of Academic Affairs, School of Dentistry
University of California, San Francisco

Phil Brown, PhD
University Distinguished Professor of Sociology and Health Sciences
Northeastern University

Adelita G. Cantu, PhD, RN
Associate Professor
Alliance of Nurses for Healthy Environments

Courtney Carignan, PhD
Assistant Professor
Department of Food Science and Human Nutrition and Department of Pharmacology and Toxicology
Michigan State University

ED_004886_00001520-00002

Jeanne A. Conry, MD, PhD
President, Environmental Health Leadership Foundation
Past President, The American Congress of Obstetricians and Gynecologists

Carl F. Cranor, PhD
Distinguished Professor of Philosophy
Faculty Member, Environmental Toxicology Graduate Program
Department of Philosophy
University of California, Riverside

Adam M. Finkel, ScD, CIH
Clinical Professor of Environmental Health Sciences, University of Michigan School of Public Health
Executive Director, Penn Program on Regulation, University of Pennsylvania Law School

Susan J. Fisher, PhD
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Robert M. Gould, MD
Associate Adjunct Professor
Program on Reproductive Health and the Environment
University of California, San Francisco
Past President, Physicians for Social Responsibility

Alycia Halladay, PhD
Chief Science Officer
Autism Science Foundation

Jean-Marie Kauth, PhD
Associate Professor
Benedictine University

Detlef R.U. Knappe, PhD
Professor
Department of Civil, Construction, and Environmental Engineering
North Carolina State University

Erica Koustas, PhD
Scientific Consultant
Program on Reproductive Health and the Environment
University of California, San Francisco

Hal C. Lawrence, III, MD, FACOG*
Executive Vice President and CEO
American Congress of Obstetricians and Gynecologists

ED_004886_00001520-00003

Heather Patisaul, PhD
Professor, Biological Sciences
Center for Human Health and the Environment, WM Keck Center for Behavioral Biology
North Carolina State University

Joshua F. Robinson, PhD
Assistant Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Ted Schettler MD, MPH
Science Director
Science and Environmental Health Network

Patrice Sutton, MPH
Research Scientist
Program on Reproductive Health and the Environment
University of California, San Francisco

Aolin Wang, PhD
Postdoctoral Scholar
Program on Reproductive Health and the Environment
University of California, San Francisco

Nsedu Obot Witherspoon, MPH
Executive Director
Children's Environmental Health Network

Lauren Zajac, MD, MPH
Assistant Professor
Department of Environmental Medicine and Public Health
Icahn School of Medicine at Mount Sinai

Marya G. Zlatnik, MD, MMS
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

*indicates organizational support

ED_004886_00001520-00004

DETAILED COMMENTS

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**

Overall, we strongly commend the EPA for its efforts to utilize a systematic and transparent method of research synthesis to reach a concise, strength of evidence conclusion about the human health hazard resulting from exposures to these ten chemicals. Efforts to integrate systematic review methods, including the explicit development of search terms, strategies, and inclusion/exclusion criteria beforehand, is relatively new in EPA's chemical assessment and as such, we applaud the EPA for this and its general improvements in its hazard assessment methodology. These scoping documents generally provide an important infrastructure for outlining EPA's screening approach for identifying relevant references and to document decisions made in the process of identifying the body of scientific literature that will be evaluated in the chemical assessments.

To improve on this document and advance EPA's uptake of systematic review methods of research synthesis, we identify the following opportunities for improvement:

EPA should not exclude studies based on language. EPA's search strategy is limited to English-only studies. The exclusive reliance on English-language studies may not represent the entire body of available evidence, and studies have suggested that language bias might lead to erroneous conclusions.[2] Furthermore, when considering the inclusion or update of an existing systematic review, studies have found that language-inclusive systematic reviews (including studies in languages other than English) were of the highest quality, compared with other types of reviews.[3] Online translation tools are readily available to allow screeners to quickly evaluate study abstracts for relevance, and therefore we recommend EPA to incorporate non-English language studies in their screening and not simply exclude these potentially relevant papers.

EPA should provide exclusion reasons for off topic citations. In the Bibliography Supplemental File for the Scope Documents, EPA has provided lists of bibliographic citations that were identified and screened from the initial literature search and the initial categorization of whether citations were *on topic* or *off topic.* We recommend EPA additionally provide exclusion reasons that were used to come to the conclusion that each citation was *off topic*, as this is a standard recommendation to fulfill transparency in documenting and reporting all decisions made in the study selection process.[4] This is particularly important as EPA has proposed to do its screening in Distiller, proprietary software that presumably will

---

[2] Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., Mierzwinski-Urban, M., Clifford, T., Hutton, B. and Rabb, D., 2012. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *International journal of technology assessment in health care*, *28*(2), pp.138-144.

[3] Moher, D., Pham, B., Lawson, M.L. and Klassen, T.P., 2003. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess*, *7*(41), pp.1-90.

[4] McDonagh M et a. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Agency for Healthcare Research and Quality. 2008.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J. and Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, *6*(7), p.e1000100.

Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. March 2011. Available from: http://handbook-5-1.cochrane.org/chapter_7/7_2_5_selecting_excluded_studies.htm

ED_004886_00001520-00005

not be made publicly available, raising concerns regarding the transparency and reproducibility of this screening step.

EPA should consider other tools for systematic review. EPA has also proposed to extract data results in the DRAGON software. We strongly encourage EPA to also consider other potential software tools that have been developed and actively incorporated into the process of systematic review, such as Swift Reviewer,[5] Active Screener,[6] HAWC (Health Assessment Workplace Collaborative).[7] These tools will help to ensure consistent and transparent execution and presentation of reviews and increase transparency of EPA assessment. Furthermore, we urge EPA to work with the National Toxicology Program and other organizations involved in these efforts in an ongoing basis to develop these and other open source tools to train scientists in their use. We believe that such infrastructure development will be critical to increasing the efficiency of chemical assessments and to expedite uptake of systematic reviews in environmental health.

EPA should have two independent reviewers for screening steps. EPA has outlined its process for screening title and abstracts of papers as having a single reviewer reviewing papers to determine whether the study is *on-topic* or *off-topic*. As part of this process, a senior-level technical expert in the topic area of interest reviewed the appropriateness of the assigned tags for "the first batch of studies" and provided feedback to the screener. Senior-level technicians also provided feedback and guidance on specific references to the individual screeners as needed during the screening and tagging process. From the description of this process, it does not appear that two independent reviewers screen all titles and abstracts for potential inclusion. Using two independent reviewers is a standard approach in systematic reviews and therefore we strongly recommend that EPA include a second independent reviewer within this process to ensure that **all** studies are screened by two reviewers at each step (title and abstract and full text).

EPA should clearly document decisions related to the identification and search. For example, it was unclear how many studies were included in the first batch of studies reviewed by the senior-level technician—these decisions should be clearly specified beforehand as to the number (or percent) that will be reviewed by this independent reviewer. Furthermore, it is unclear how many studies the senior-level technical experts are reviewing generally as to their additional feedback and guidance to individual screeners. This should be more clearly stated and described beforehand in these protocols.
We recommend EPA broaden the set of studies that are initially screened in the first batch to ensure consistency across reviewers and demonstrated understanding of protocol instructions by all reviewers before moving on to screening the remaining records. It is stated in the Gray Literature Search Results that individual screeners would screen and tag 10 references that would be then independently reviewed by the senior-level technical expert. However, this does not seem to be an adequate number of studies as it is a small number relative to the expected number of records that will ultimately be screened.

EPA should clarify how it will handle discrepancies in the inclusion/exclusion and tagging process. As it is stated in its current protocol, it appears that the senior-level technical expert has the final say in

---

[5] Howard, B.E., Phillips, J., Miller, K., Tandon, A., Mav, D., Shah, M.R., Holmgren, S., Pelch, K.E., Walker, V., Rooney, A.A. and Macleod, M., 2016. SWIFT-Review: a text-mining workbench for systematic review. *Systematic reviews*, *5*(1), p.87.
[6] https://www.sciome.com/swift-activescreener/
[7] https://hawcproject.org/about/

determining the final inclusion/exclusion decision and tagging, for the subset of studies they evaluate. However, this should be clarified and we also highly recommend that a third party reviewer be incorporated as an arbiter for these decisions if consensus cannot be reached between the two reviewers, as is typically standard in systematic reviews.

EPA should stratify its exclusion criteria separately at the title and abstract and full-text screening steps. It is likely that title and abstracts of papers would not contain sufficient detail to evaluate all exclusion criteria—many of these would likely only be identified in the full-text of the paper. To increase the efficiency of the screening process, it would help to create a subset of exclusion criteria most relevant when screening the title and abstracts of records versus the larger set of exclusion criteria relevant to screening the full text of records.

EPA should clearly outline the process for handling anticipated overlap with literature relevant to multiple topics. EPA should describe how this will be addressed by the screeners and whether the same reviewer will be responsible for screening papers with inclusion/exclusion criteria across multiple topics or whether different reviewers are responsible only for screening studies for one particular topic. Additional details in regards to the process by which this screening will occur would be helpful. Given the breadth of each assessment (searching literature related to fate, engineering, exposure, human health, and environmental hazard) and the complexity of the screening process (tagging *on-topic* and *off-topic* literature and using additional sub-categories or sub-tags to allow for additional categorization), there appears to be the potential for individual papers to fall into different topic categories and have many different tags and sub-tags applied to indicate their relevance. However, it is unclear how this will be organized in the screening phase. Search strategies and inclusion/exclusion criteria appear to have been developed specifically for each literature topic and the potential overlap of literature relevant to multiple topics is not addressed.

EPA should explicitly include stopping rules for when the list of relevant studies will be considered final. There is no discussion of stopping dates or the process of updating the literature search to search for newer studies. Newer scientific studies will inevitably continue to appear in scientific journals and it will be impossible to continually attempt to include all these studies in a chemical assessment. To meet the deadlines as mandated by the Lautenberg Amendments, EPA should state clear stopping rules in the form of deadlines or criteria for when the body of included relevant studies will be finalized for the purposes of the chemicals assessment.

EPA should ensure gray literature search results are adequately screened. EPA's gray literature search strategy proposes to utilize Google's API to develop custom searches and return the first 100 results, sorted by predicted relevancy so that the results likely to be most relevant are screened first. It is unclear why this number is limited to only the first 100 and whether there was an empirical reason for why this particular number was selected. We recommend that EPA ensure that an adequate number of search results are screened, in particular considering that the gray literature can contribute potentially important information relevant to toxicity, mode of action, exposure, fate and transport, engineering or occupational exposure, or existence of publication bias.

EPA should consider "snowball searching," where the citations of included (i.e., *on-topic*) references are searched as well as using databases such as Web of Science to search for references that cite the included citations. EPA states that it plans on assessing the specificity and efficiency of the literature searches, through comparison of references either cited in existing problem formulation and risk assessment documents, in the public use documents and supporting life cycle diagrams, and comparison

of the references cited in review articles. Snowball searching will contribute to the evaluation of the specificity and efficiency of its literature searches, and also help to identify newer relevant studies that could potentially be included that have not yet been indexed in main databases such as PubMed.

EPA should incorporate appropriate tools for updating and evaluating systematic reviews in their chemical assessments. Garner et al., as part of efforts by a Cochrane Collaboration panel for updating guidance for systematic reviews, published guidance in 2016 for determining when it is appropriate to update a systematic review and outlining the steps for performing the update.[8] We have attached this guidance as an Appendix to these comments. EPA should evaluate the Cochrane tool's applicability to environmental chemicals given that Cochrane systematic reviews are geared towards reviews of clinical intervention evidence, so these tools may require updating and tailoring for an application to environmental health data.

A recent NAS report[9] recommends EPA should develop policies and procedures to allow the agency to identify, use and update existing systematic reviews. The committee also noted that it was important that the existing systematic review's study question directly addresses EPA's topic of interest and that the methods are critically evaluated before the systematic review is used and updated. EPA should ensure that only the highest quality systematic reviews be considered appropriate for use. It will be critical for EPA to develop tools to assist with the process of evaluating existing systematic reviews, particularly as this field continues to rapidly expand and more systematic reviews relevant to environmental health questions are published in the scientific literature, potentially of variable quality.

One tool which might be helpful for evaluating the risk of bias in systematic reviews is the ROBIS tool, which the NAS committee utilized in their report.[10] This tool was developed using rigorous methodology and can be applied for evaluating internal validity of systematic reviews in conjunction with other available tools to critically appraise and assess their quality. Of particular note is the strong emphasis on the recommendation that tools such as ROBIS should not be used to generate a composite quality score, as it has been well-documented that scoring can lead to bias in evaluation of the studies.[11] As such, the ROBIS tool presents several options for visually and graphically presenting results from risk of bias assessments based on individual domains or the overall rating, enabling reviewers to highlight particular areas of concern or reviews that are most relevant to the target question of interest.

Another tool which may be helpful in this process is the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), used by authors of systematic reviews to improve the reporting

---

[8] Garner P, Hopewell S, Chandler J, MacLehose H, Akl EA, Beyene J, Chang S, Churchill R, Dearness K, Guyatt G, Lefebvre C. When and how to update systematic reviews: consensus and checklist. bmj. 2016 Jul 20;354:i3507.

[9] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[10] Id.

[11] Jüni, P., Witschi, A., Bloch, R. and Egger, M., 1999. The hazards of scoring the quality of clinical trials for meta-analysis. *Jama*, *282*(11), pp.1054-1060.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

ED_004886_00001520-00008

of elements relevant to the systematic review and meta-analyses.[12] Increasingly, scientific journals are requiring the inclusion of checklists such as PRISMA with the submission of systematic reviews considered for publication. Although this tool is used to evaluate study reporting, and is not an assessment instrument to gauge the quality of a systematic review, it can still provide a useful framework to identify reported components of an existing systematic review in the process of evaluating quality or to identify missing components requiring follow-up with study authors to obtain additional information.

Furthermore, we strongly encourage EPA to evaluate the potential for financial conflicts of interest as an element in their study design. This is currently included as a consideration in evaluation risks of bias in some frameworks, such as the Navigation Guide,[13] and extracted for consideration as an additional domain in other frameworks, such as that developed by the National Toxicology Program's (NTP) Office of Health Assessment and Translation (OHAT).[14] The Cochrane Collaboration's risk of bias tool does not currently include a specific domain for bias related to study funding source, but this is an area of active discussion among its members.[15] The Cochrane Collaboration has recognized the importance of identifying study funding source, which has been empirically shown to be associated with biases.[16] A recent report from the NRC recommended that the U.S. EPA consider funding sources in their risk of bias assessment conducted for systematic reviews.[17]

We also strongly recommend EPA identify tools that may potentially not be appropriate for human health chemical assessments. Many tools are currently being developed for evaluating risk of bias, quality, and strength of evidence for individual studies as well as for systematic reviews. It is critical that EPA evaluate tools developed in other fields that may be relevant, such as for clinical or preclinical animal or human studies, as these tools could potentially be modified for an application to questions of environmental health relevance. However, these tools should be applied with caution—due to the

---

[12] http://www.prisma-statement.org/; Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews. 2015 Jan 1;4(1):1.

[13] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

[14] NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[15] Bero LA. Why the Cochrane risk of bias tool should include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000075.*

Sterne JA. Why the Cochrane risk of bias tool should not include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000076.*

[16] Krauth D, Anglemyer A, Philipps R, Bero L. Nonindustry-sponsored preclinical studies on statins yield greater efficacy estimates than industry-sponsored studies: a meta-analysis. *PLoS Biol. 2014 Jan; 12(1):e1001770*

Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev. 2012 Dec 12; 12():MR000033.*

[17] NRC (National Research Council). Review of EPA's Integrated Risk Information System (IRIS) Process. Washington, DC:National Academies Press. 2014.

ED_004886_00001520-00009

differences in the types of evidence under evaluation a direct application to a difference evidence base than intended could lead to biased and erroneous conclusions.[18]

2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**

In general, EPA is proposing to consider three populations for exposure assessment: 1) Occupational users and non-users; 2) consumers and bystanders; and 3) general population. We strongly recommend that EPA calculate the aggregate exposures within and across these populations-- risk will be underestimated if it does not include these real-world exposures. Exposures within a population should also be aggregated (rather than considered in isolation) in order to estimate the general population's actual exposure to the chemical—for example, through exposures from food, water and air.

Further, as shown in the Figure below, exposures must also be aggregated across populations. Consumers and workers are part of the general population – that is, since workers and consumers also eat food and drink water, they will have the same exposures as the general population, in addition to the anticipated exposures on-the-job or from consumer products. Some workers will also be consumer product users, so they have the potential to face general, consumer product, and on-the-job exposures. These specific exposure scenarios must be accounted for in EPA's exposure estimation to ensure that such individual exposures are adequately considered and integrated into the risk assessment.

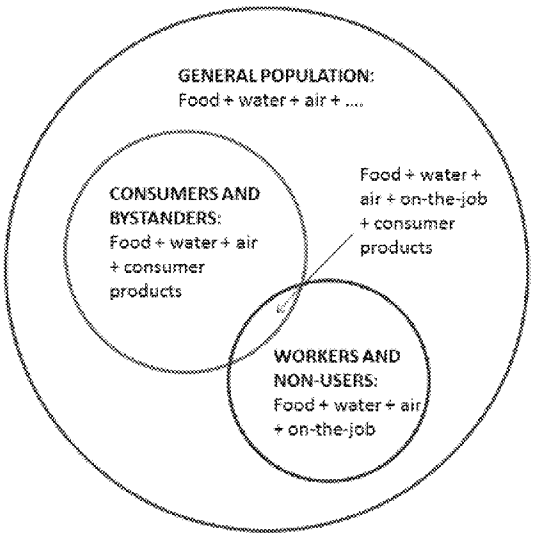


Figure: EPA must assess aggregate exposures within and across all the populations for accurate exposure assessment.

In the Introduction section of the chemical Scope documents, EPA states that it "may consider background exposures from legacy use, associated disposal, and legacy disposal as part of an assessment of aggregate exposure or as a tool to evaluate the risk of exposures resulting from non-

[18] Nachman, K.E., Lam, J., Schinasi, L.H., Smith, T.C., Feingold, B.J. and Casey, J.A., 2017. O'Connor et al. systematic review regarding animal feeding operations and public health: critical flaws may compromise conclusions. *Systematic Reviews*, *6*(1), p.179.

legacy uses."[19] This falls short of the analysis required under Lautenberg TSCA. It is critical that EPA consider ongoing exposures from legacy uses and disposal, and includes these as part of the aggregate exposure assessment. Asbestos and HBCD are two examples of this, as they have enormous volumes in place in buildings and existing infrastructure. The Healthy Building Network estimates there are 66 million- 132 million pounds (30,000-60,000 metric tons) of HBCD in insulation in existing buildings[20] — these reservoirs in-place are and will continue to be critical sources of ongoing exposures. HBCD was also used in cars and furniture, which are long-lived consumer items that will continue to contribute to ongoing exposures for years to come.

Another example is 1,4-dioxane, which was historically used as a chemical stabilizer for chlorinated solvents. Many groundwater aquifers are contaminated with 1,4-dioxane, and the extent of legacy contamination of groundwater is likely underestimated.[21] Also, 1,4-dioxane occurs in a wide variety of products including personal care products, detergents, waxes, and antifreeze, and 1,4-dioxane is a by-product in manufacturing processes involving ethylene oxide, such as the production of polyethylene terephthalate (PET), polyester, and surfactants. The use and disposal of 1,4-dioxane has led to past environmental contamination which contributes to on-going exposures.[22] The physical and chemical properties of 1,4-dioxane render it a persistent and highly mobile water contaminant: it is highly miscible in water.[23] Exposures via drinking water are documented back to the 1980s and continue today.[24] Results from EPA's Third Unregulated Contaminant Monitoring Rule (UCMR3) highlight that over 13% of 4,905 public drinking water systems serving >10,000 people had concentrations of 1,4-dioxane above the EPA Reference Concentration of 0.35 ppb 1,4-dioxane.[25] Furthermore, the UCMR3 results do not capture exposures in communities served by small public drinking water systems serving <10,000 people.[26] Approximately 27% of the US population is served by small public drinking water systems.[27] Thus, it will be critical for EPA to consider the population's current exposure to 1,4-dioxane via sources like drinking water as part of their assessment for health risks.

---

[19] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 12

[20] Safer Chemicals, Healthy Families et al. Comments to the U.S. Environmental Protection Agency (EPA) on the Scope of its Risk Evaluation for the TSCA Work Plan Chemicals: CYCLIC ALIPHATIC BROMIDE CLUSTER or HEXABROMOCYCLODODECANE (HBCD). March 15, 2017. https://healthybuilding.net/uploads/files/saferchemicals-hbcd.pdf

[21] David T. Adamson, Shaily Mahendra, Kenneth L. Walker, Jr., Sharon R. Rauch, Shayak Sengupta, and Charles J. Newell. A Multisite Survey To Identify the Scale of the 1,4-Dioxane Problem at Contaminated Groundwater Sites. *Environmental Science & Technology Letters* 2014 1 (5), 254-258 DOI: 10.1021/ez500092u

[22] Agency for Toxic Substances and Disease Registry. TOXICOLOGICAL PROFILE FOR 1,4-DIOXANE. Atlanta, GA; 2012.

[23] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

[24] Dietrich, A.M., D.S. Millington, and R.F. Christman. 1983. Specific identification of organic pollutants in Haw River water using gas chromatography/mass spectrometry. WRRI Report No. 206. North Carolina State University, Raleigh, NC.

[25] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

U.S. Environmental Protection Agency. Unregulated Contaminant Monitoring Rule 74 FR 51850, EPA-HQ-OW-2007-1189 FRL-8963-6. Washington, D.C.; 2009. https://federalregister.gov/a/E9-24287.s

[26] Knappe, D.R.U., Lopez-Velandia, C., Hopkins, Z., & Sun, M. (2016). Occurrence of 1,4-Dioxane in the Cape Fear River Watershed and Effectiveness of Water Treatment Options for 1,4-Dioxane Control. NC Water Resources Research Institute.

[27] EPA. 2017. Small Drinking Water Systems Research https://www.epa.gov/water-research/small-drinking-water-systems-research-0 (accessed September 17, 2017).

ED_004886_00001520-00011

When a chemical is present in products or media as a contaminant/ by-product, EPA needs to include and assess these exposures. We strongly recommend against ignoring or discounting these potential exposures routes. For example, EPA proposes to exclude from consideration conditions of use of 1,4-dioxane when it is present as contaminant in a wide variety of items, including household detergents, cosmetics/ toiletries, and foods.[28] This exclusion is not scientifically justified. Cosmetics and personal care products have the potential to contribute significantly to exposures, since people are applying them directly to their bodies, often multiple times per day, every day.

Finally, in the exposure assessments for methylene chloride, N-methylpyrrolidone and trichloroethylene, EPA is proposing to exclude uses it already assessed.[29] We agree that EPA does not need to re-assess these uses; these evaluations have been completed and finalized. However, unless and until such uses are banned, the exposures from these uses continue. Therefore, the new risk evaluations need to consider the contributions of these uses to exposures by using the exposure values from the previous assessments.

For the occupational exposure analysis plan, EPA states it will "Consider and incorporate applicable engineering controls and/or personal protective equipment into exposure scenarios."[30] However, these are not realistic assumptions nor are they appropriate for public health protection. EPA's own research shows that the primary factors influencing whether a user understands label information are the users' literacy and numeracy, which frequently correlate with the users' education and income.[31] Therefore, people with less education, lower income, and less advanced literary skills will be the most likely to not understand label instructions. These individuals already disproportionately bear the burden of exposures to multiple environmental hazards and the resulting health impacts; thereby placing further burden on this already stressed susceptible subpopulation. Further, appropriate personal protective equipment (PPE) for workers is often not provided by employers, or may not be fitted or working properly. When evaluating occupational exposures, EPA needs to take into consideration all potential and feasible routes of exposure, and should not exclude exposure routes based on assumptions of PPE and/ or exposure controls in place. These controls are not guaranteed and may change in the future, so to assume zero exposure via these routes would be inappropriate and a failure to adequately ensure health protections, especially for susceptible sub-populations as required by the Lautenberg TSCA.

In summary, EPA needs to account for all the sources of exposure or it will underestimate risk for all 10 chemicals. When analyzing aggregate exposures, "sentinel exposure" may be considered simultaneously, where appropriate. However, these are not mutually exclusive and EPA should not incorporate sentinel to the exclusion of aggregate.

3. EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations

---

[28] US EPA (2017). Scope of the Risk Evaluation for 1,4-dioxane. Pg. 21

[29] US EPA (2017). Scope of the Risk Evaluation for Methylene Chloride (Dichloromethane, DCM). Pg. 30
US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 19-20
US EPA (2017). Scope of the Risk Evaluation for Trichloroethylene. Pg. 27

[30] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 45

[31] US Environmental Protection Agency (EPA). 2016. "The Effectiveness of Labeling on Hazardous Chemicals and Other Products." Office of Chemical Safety and Pollution Prevention. RIN 2079-AK07.

ED_004886_00001520-00012

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

In general, EPA proposes to consider workers and occupational non-users, consumer and by-standers, and other groups within the general population in proximity to conditions of use as sub-populations who experience greater exposures. In particular, EPA has appropriately identified people who live or work near manufacturing, processing, distribution, use or disposal sites as facing greater exposures. Such communities are often low income and/ or people of color, exposed to a disproportionate share of pollution, environmental hazards, social and economic stressors. Multiple exposures to chemical and non-chemical stressors collectively increase the risk of harm, combined with synergistic effects with other health stressors in their daily lives such as limited access to quality health care.[32,33]

EPA's risk evaluation needs to fully account for the reality of cumulative exposures, as recommended by the NAS in their Phthalates and Cumulative Risk report.[34] As described below, EPA can use "default values" to account for cumulative exposures.

In regards to greater susceptibility, EPA's considerations for addressing susceptibility vary considerably across the 10 chemicals. EPA should apply a consistent approach to addressing susceptibility across the 10 chemicals. The following are well-known factors that increase biologic sensitivity or reduce resilience to exposures,[35,36] and these as well as other relevant factors should be standard considerations for all 10 chemicals to identify susceptible sub-populations:

Intrinsic/ endogenous factors

- Genetic polymorphisms/ genetics/ genetic makeup
- Health status/ nutritional status/ disease status/ pre-existing conditions
- Prenatal lifestage
- Age

Extrinsic factors

- Multiple exposures/ co-exposures
- Race/ ethnicity
- Socioeconomic status (SES)

For example, the prenatal lifestage is the most sensitive to developmental and reproductive toxicants, and women of child-bearing age should be considered as a susceptible sub-population for any chemicals with such hazards. Women of reproductive age are not specifically identified as a potential susceptible

---

[32] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[33] Vesterinen HM, Morello-Frosch R, Sen S, Zeise L, Woodruff TJ. Cumulative effects of prenatal-exposure to exogenous chemicals and psychosocial stress on fetal growth: Systematic-review of the human and animal evidence. Meliker J, editor. PLoS One. 2017 Jul 12;12(7):e0176331.

[34] National Research Council. Committee on the Health Risks of Phthalates, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies. 2008. Phthalates and cumulative risk assessment: the task ahead. Washington, D.C.: National Academies Press.

[35] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[36] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

ED_004886_00001520-00013

sub-population for pigment violet 29, TCE, NMP, PERC, or HBCD, even though EPA will consider reproductive and developmental toxicity hazards for these chemicals.

As discussed below, science-based defaults should be used to account for these and other susceptibilities, unless there is there is chemical-specific data available to support increasing or decreasing the default.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**

EPA cites existing IRIS assessments for five chemicals; because these are EPA's own assessments which have gone through the Agency's peer-review process, and in some cases NAS review, EPA can rely on these existing finalized, authoritative assessments for hazard identification.

Moving forward, a weight of evidence evaluation is required by law, which EPA defines as:

> "Weight of scientific evidence means a systematic review method, applied in a manner suited to the nature of the evidence or decision, that uses a pre-established protocol to comprehensively, objectively, transparently, and consistently, identify and evaluate each stream of evidence, including strengths, limitations, and relevance of each study and to integrate evidence as necessary and appropriate based upon strengths, limitations, and relevance."[37]

Therefore, EPA should use a systematic review process for evaluating scientific information for chemicals that do not have an IRIS assessment and for any additional studies that will be considered for the chemicals that have IRIS assessments.

For the scoping document, EPA should include all hazards identified in the literature, and not make decisions about their relevance to the risk evaluation until a systematic review has been completed. For a number of chemicals, EPA has inappropriately drawn conclusions about hazards prior to the completion of a systematic review. Some examples are given in the table below where EPA concludes that HBCD, NMP and pigment violet 29 are not genotoxic based on previous assessments and without conducting a systematic review.

Chemical	Example Text from EPA Scoping Document
HBCD	"Available data suggest that HBCD is not genotoxic. Existing assessments have also concluded, based on genotoxicity information and a limited lifetime study, that HBCD is not carcinogenic (NICNAS, 2012; EINECS, 2008; TemaNord, 2008; OECD, 2007). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity or cancer hazards in the risk evaluation of HBCD at this time."[38]
NMP	"NMP is not mutagenic, based on results from bacterial and mammalian *in vitro* tests and *in vivo* systems and is not considered to be carcinogenic (RIVM, 2013; OECD,

[37] 82 Fed. Reg. 138, 33748
[38] US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg.36

	2007; WHO, 2001). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity and cancer hazards in the NMP risk evaluation."[39]
Pigment violet 29	"Testing for carcinogenicity of Pigment Violet 29 has not been conducted. However, negative genotoxicity results, structure-activity considerations and the expectation of negligible absorption and uptake of Pigment Violet 29 (based on very low solubility), indicate carcinogenicity of Pigment Violet 29 is unlikely. Unless new information indicates otherwise, EPA does not expect to conduct additional, in-depth analyses of genotoxicity and cancer hazards in the risk evaluation of Pigment Violet 29."[40]

The National Academies recently released a report with recommendations on implementation of systematic review for EPA's chemical evaluations (which we will refer to as the 'NAS Systematic Review report' for simplicity).[41] First, they recommend that EPA should develop policies and procedures that allow the agency to use and update existing systematic reviews, since the committee concluded that could potentially save time and resources. EPA should conduct a review to determine whether there are existing systematic reviews on the topic of interest and if there is, EPA should evaluate it to determine if it is high-quality. The NAS recommends that EPA should build on existing high-quality reviews to incorporate new studies and use the updated systematic review as a basis for its assessment. The assessments cited by EPA to support the hazard identification claims are not systematic reviews; even if they were, EPA should evaluate them for quality before relying on their conclusions.

Second, it is very likely that additional studies have been published since the assessments EPA cites were completed. EPA should develop criteria to evaluate the internal validity (risk of bias) of individual studies, utilizing existing tools that have been developed and empirically demonstrated on environmental health studies such as the Navigation Guide or the OHAT approach.[42] We also recommend that EPA not using a scoring system to evaluate study quality. Specifically, we note that empirically validated approaches in the clinical sciences such as Cochrane discourage using a numerical scale scoring approach for evaluating study quality because calculating a score requires choosing a weighting scheme for each component, which generally is nearly impossible to justify.[43] Furthermore, a study might be well designed to eliminate bias, but because the study failed to report details in the publication under review, it will receive a low score--most available scoring systems include a mix of risk of bias and reporting biases which is inappropriate. Additionally, quality scores have been shown to be invalid for assessing risk of bias in clinical research.[44] The current standard in evaluation of both clinical

[39] US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 36

[40] US EPA (2017). Scope of the Risk Evaluation for Pigment Violet 29. Pg.29

[41] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[42] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[43] Juni, P., A. Witschi, R. Bloch, and M. Egger. 1999. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 282(11):1054-1060.

[44] Id.

ED_004886_00001520-00015

and environmental health research calls for reporting each component of the assessment tool separately and not calculating an overall numeric score.[45]

Data generated by alternative test methods (such as high-throughput screening methods) are not different than any other type of *in vitro* or cell-based assay data that would be considered in a systematic review. These kinds of assays provide mechanistic data, and the NAS Systematic Review report explicitly considered how mechanistic data could be utilized in a systematic review for evidence integration. The committee came to two conclusions. First, the same protocol for evaluating relevance and study quality must be used with mechanistic data as for any other study. For example, in the report's case study on phthalates, the committee was not able to integrate results from high-throughput assays because the cell lines used were of unknown relevance to the *in vivo* mechanism of phthalate toxicity.[46] Second, the foundation of the hazard classification in a systematic review is the animal and human data, with the mechanistic data playing a supporting role. If mechanistic data is relevant, it can be used to upgrade a hazard classification, or increase the confidence of a finding made based on evaluation of animal and human data. A hazard classification is never made based on high-throughput or other kinds of mechanistic data alone.[47]

5. For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.

Defaults

We strongly support the use of health protective defaults to incorporate factors that reflect the range of variability and susceptibility in the population to ensure risks are not underestimated. The importance of using protective science-based defaults was highlighted by the NAS in 2009.[48] The default should be used for factors that are known to influence risk unless there is chemical-specific data that support increasing or decreasing it; when there is inadequate information to quantitatively assess inter- or intra-species differences for a specific chemical, the defaults should be used. For example, EPA's defaults should include:

- Inter-human variability, general
- Inter-human susceptibility to carcinogens, adult
- Inter-human susceptibility to carcinogens, early life (including prenatal)
- Inter-human susceptibility to non-carcinogens, early life (including prenatal)
- Animal findings are relevant to humans
- Findings from one route of exposure are considered representative unless data show otherwise

---

[45] Higgins, J.P.T., and S. Green, eds. 2008. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

[46] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017. Pg. 78

[47] Id. Pp. 158-9

[48] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch 4-6

ED_004886_00001520-00016

EPA has relied on standard default values ("uncertainty" or "safety" factors) that have been applied across the board to various chemicals and health outcomes. But newer science demonstrates that EPA's typical safety factor of 10 is insufficient to account for variability due to life stage, genetics, underlying disease status, and external stressors that may be due to poverty or other difficult life conditions.

For cancer, the NAS recommended that EPA include a factor to account for human variability in response to carcinogens, as EPA's current approach inaccurately assumes that there is no variability in response. They found that a factor of 25- to 50- may account for the variability between the median individual and those with more extreme responses, and recommended 25 as a reasonable default value.[49]

Similarly, EPA should increase or add factors that address cancer and non-cancer susceptibility during early life stages. While EPA does account for increased susceptibility to genotoxicants, it does not include the prenatal period or chemicals that can influence cancer through other mechanisms. California EPA's guidance incorporates factors to account for increased susceptibility for exposures that occur prenatally for carcinogens, non-mutagenic carcinogenic agents and non-carcinogens. Their literature review on differential susceptibility to carcinogens and non-carcinogens based on age and life stage derived age adjustment values for carcinogens which include the prenatal period[50] and increased the default intraspecies uncertainty factors for non-carcinogens to 30 and 100 for specific endpoints such as asthma or neurotoxicity.[51] At a minimum, EPA should use Cal EPA's age adjustment values and intraspecies uncertainty factors for incorporating age/early life susceptibility.

In general, developmental life stages, including the fetus, infancy, and childhood, are more vulnerable to chemical exposure and toxicity. However, typical EPA age-dependent adjustment factors account for other life stages but NOT fetal exposures. Recent studies have demonstrated differential expression and activity of metabolic enzymes such as Cytochrome P450 in fetal versus adult tissue, indicating potential lifestage-dependent variability in metabolic capabilities and greater vulnerability during fetal development not accounted for in current risk assessment practices.[52] This is a critical point to address, as disruptions during fetal development have implications for health and disease in adulthood. EPA should evaluate this rich body of literature to identify the most up-to-date scientific knowledge regarding human variability and susceptibility and incorporate these scientifically-based default values in their assessments unless there are chemical-specific data supporting departing from the defaults. California EPA also developed child-specific risk values for chemicals (e.g., atrazine, lead, nickel, manganese, heptachlor) that specifically address routes of exposure and differences in susceptibility

---

[49] Id. Pg. 168

[50] California EPA 2009. Cal EPA 2009. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document for Cancer Potency Factors: Methodologies for derivation, listing of available values, and adjustments to allow for early life stage exposures. http://oehha.ca.gov/media/downloads/crnr/tsdcancerpotency.pdf

[51] Cal EPA 2008. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document For the Derivation of Noncancer Reference Exposure Levels http://oehha.ca.gov/media/downloads/crnr/noncancertsdfinal.pdf

[52] Sadler, N.C., Nandhikonda, P., Webb-Robertson, B.J., Ansong, C., Anderson, L.N., Smith, J.N., Corley, R.A. and Wright, A.T., 2016. Hepatic cytochrome P450 activity, abundance, and expression throughout human development. *Drug Metabolism and Disposition*, *44*(7), pp.984-991.

ED_004886_00001520-00017

unique to children compared to adults.[53] EPA should review these evaluations and incorporate these values as appropriate. Furthermore, a default guidance principle should be that animal findings are relevant to humans unless there is sufficient and compelling information to support otherwise.

Risk estimates

EPA should not use MOE (margin of exposure) as an analysis method in the risk evaluation process moving forward. MOE is not an estimate of risk—it is a single number that is a version of the "bright line" approach like the Reference Dose (or Reference Concentration for inhalation doses). MOE is calculated by dividing the point of departure (e.g., LOAELs, NOAELs or BMDLs) by estimated exposure values, and this 'bright line' approach does not provide information about the magnitude of the risks above, at, or below this line. Further, it implies that there is a "safe" level of exposure below which no harm will occur. While this may be true for a select few chemicals, the NAS Science and Decisions report recognizes that this is not a valid assumption for all chemicals and has recommended moving away from such "bright line" approaches which do not establish risk estimates across the full range of exposures.[54] Additionally, the MOE will not provide the necessary information for future analysis of risks and benefits that will be critical for decision-making on these chemicals.[55] We recommend that EPA utilize available analytical methods such as PODs based on a BMD to develop quantified estimates of risk.

EPA appropriately states that a dose-response assessment will be conducted for all identified human health hazard endpoints. PODs should also be developed for every endpoint unless the data are insufficient to develop a model. For calculating cancer or non-cancer risks, we recommend always using a point of departure (POD) of a benchmark dose (BMD) at 1%. The POD should be based on a BMD calculation, not the NOAEL/LOAEL, unless the data are insufficient to model. EPA already recognizes the features that make BMDs superior: BMDs account for the shape of the dose–response function; are independent of study design, such as the space between dosing; and are comparable across chemicals.[56]

Historically, for carcinogens that are direct mutagens or are associated with large human body burdens, EPA has assumed there is no threshold of effect. But the NAS Science and Decisions report highlights the science indicating that this linear presumption with no threshold is appropriate for the calculation of both cancer and non-cancer risks, and regardless of whether a carcinogen is a mutagen. For example, dose-response relationships can be linear at low dose when exposures contribute to an existing disease process, add to background processes and/ or exposures, and interact with interindividual variability or susceptibility.[57] Science and Decisions recommends harmonizing cancer and non-cancer risk assessment approaches. Therefore, for calculating non-mutagen cancer or non-cancer risks based on a POD, EPA should use the same approach as for mutagens, which assumes a straight line from the POD. In fact, a

---

[53] California Environmental Protection Agency. Office of Environmental Health Hazard Assessment (OEHHA). Child-Specific Reference Doses (chRDs) Finalized to Date. Available from: http://oehha.ca.gov/risk-assessment/chrd/table-all-chrds

[54] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

[55] McGartland A, Revesz R, Axelrad DA, Dockins C, Sutton P, Woodruff TJ. Estimating the health benefits of environmental regulations. Science (80-). 2017 Aug 4;357(6350):457–8.

[56] Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, Rusyn I. 2014. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. Environmental health perspectives. 122(5).

[57] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch.5

linear relationship may actually underestimate risks for some chemicals where the dose-response curve is supra-linear.

6. Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.

Production volumes for both asbestos and pigment violet 29 have been claimed as CBI. Production volume is basic information about a chemical to which the public and scientists should have access. We urge EPA to move forward with substantiating such claims under the new TSCA.

Appendix

When and how to update systematic reviews: consensus and checklist
Garner et al.
BMJ 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

ED_004886_00001520-00020

OPEN ACCESS

When and how to update systematic reviews: consensus and checklist

Paul Garner,[1] Sally Hopewell,[2] Jackie Chandler,[3] Harriet MacLehose,[3] Elie A Akl,[5,6] Joseph Beyene,[7] Stephanie Chang,[8] Rachel Churchill,[9] Karin Dearness,[10] Gordon Guyatt,[4] Carol Lefebvre,[11] Beth Liles,[12] Rachel Marshall,[3] Laura Martínez García,[13] Chris Mavergames,[14] Mona Nasser,[15] Amir Qaseem,[16,17] Margaret Sampson,[18] Karla Soares-Weiser,[3] Yemisi Takwoingi,[19] Lehana Thabane,[4,20] Marialena Trivella,[21] Peter Tugwell,[22] Emma Welsh,[23] Ed C Wilson,[2,4] Holger J Schünemann[4,5]

For numbered affiliations see end of article.

Correspondence to: P Garner Paul.Garner@istmed.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

Accepted: 26 May 2016

Updating of systematic reviews is generally more efficient than starting all over again when new evidence emerges, but to date there has been no clear guidance on how to do this. This guidance helps authors of systematic reviews, commissioners, and editors decide when to update a systematic review, and then how to go about updating the review.

Systematic reviews synthesise relevant research around a particular question. Preparing a systematic review is time and resource consuming, and provides a snapshot of knowledge at the time of incorporation of data from studies identified during the latest search. Newly identified studies can change the conclusion of a review. If they have not been included, this threatens the validity of the review, and, at worst, means the review could mislead. For patients and other healthcare consumers, this means that care and policy development might not be fully informed by the latest research; furthermore, researchers could be misled and carry out research in areas where no further research is actually needed.[1] Thus, there are clear benefits to updating reviews, rather than duplicating the entire process as new evidence emerges or new methods develop. Indeed, there is probably added value to updating a review, because this will include taking into account comments and criticisms, and adoption of new methods in an iterative process.[2-6]

Cochrane has over 20 years of experience with preparing and updating systematic reviews, with the publication of over 6000 systematic reviews. However, Cochrane's principle of keeping all reviews up to date has not been possible, and the organisation has had to adapt: from updating when new evidence becomes available,[7] to updating every two years,[8] to updating based on need and priority.[9] This experience has shown that it is not possible, sensible, or feasible

SUMMARY POINTS

Updating systematic reviews is, in general, more efficient than starting afresh when new evidence emerges. The panel for updating guidance for systematic reviews (PUGs; comprising review authors, editors, statisticians, information specialists, related methodologists, and guideline developers) met to develop guidance for people considering updating systematic reviews. The panel proposed the following:

1. Decisions about whether and when to update a systematic review are judgments made for individual reviews at a particular time. These decisions can be made by agencies responsible for systematic review portfolios, journal editors with systematic review update services, or author teams considering embarking on an update of a review.
2. The decision needs to take into account whether the review addresses a current question, uses valid methods, and is well conducted; and whether there are new relevant methods, new studies, or new information on existing included studies. Given this information, the agency, editors, or authors need to judge whether the update will influence the review findings or credibility sufficiently to justify the effort in updating it.
3. Review authors and commissioners can use a decision framework and checklist to navigate and report these decisions with "update status" and rationale for this status. The panel noted that the incorporation of new synthesis methods (such as Grading of Recommendations Assessment, Development and Evaluation (GRADE)) is also often likely to improve the quality of the analysis and the clarity of the findings.
4. Given a decision to update, the process needs to start with an appraisal and revision of the background, question, inclusion criteria, and methods of the existing review.
5. Search strategies should be refined, taking into account changes in the question or inclusion criteria. An analysis of yield from the previous edition, in relation to databases searched, terms, and languages can make searches more specific and efficient.
6. In many instances, an update represents a new edition of the review, and authorship of the new version needs to follow criteria of the International Committee of Medical Journal Editors (ICMJE). New approaches to publishing licences could help new authors build on and re-use the previous edition while giving appropriate credit to the previous authors.

The panel also reflected on this guidance in the context of emerging technological advances in software, information retrieval, and electronic linkage and mining. With good synthesis and technology partnerships, these advances could revolutionise the efficiency of updating in the coming years.

ED_004886_00001520-00021

to continually update all reviews all the time. Other groups, including guideline developers and journal editors, adopt updating principles (as applied, for example, by the *Systematic Reviews* journal; https://systematicreviewsjournal.biomedcentral.com/).

The panel for updating guidance for systematic reviews (PUGs) group met to draw together experiences and identify a common approach. The PUGs guidance can help individuals or academic teams working outside of a commissioning agency or Cochrane, who are considering writing a systematic review for a journal or to prepare for a research project. The guidance could also help these groups decide whether their effort is worthwhile.

Panel selection and procedures

An international panel of authors, editors, clinicians, statisticians, information specialists, other methodologists, and guideline developers was invited to a two day workshop at McMaster University, Hamilton, Canada, on 26-27 June 2014, organised by Cochrane. The organising committee selected the panel (web appendix 1). The organising committee invited participants, put forward the agenda, collected background materials and literature, and drafted the structure of the report.

The purpose of the workshop was to develop a common approach to updating systematic reviews, drawing on existing strategies, research, and experience of people working in this area. The selection of participants aimed on broad representation of different groups involved in producing systematic reviews (including authors, editors, statisticians, information specialists, and other methodologists), and those using the reviews (guideline developers and clinicians). Participants within these groups were selected on their expertise and experience in updating, in previous work developing methods to assess reviews, and because some were recognised for developing approaches within organisations to manage updating strategically. We sought to identify general approaches in this area, and not be specific to Cochrane; although inevitably most of the panel were somehow engaged in Cochrane.

The workshop structure followed a series of short presentations addressing key questions on whether, when, and how to update systematic reviews. The proceedings included the management of authorship and editorial decisions, and innovative and technological approaches. A series of small group discussions followed each question, deliberating content, and forming recommendations, as well as recognising uncertainties. Large group, round table discussions deliberated further these small group developments. Recommendations were presented to an invited forum of individuals with varying levels of expertise in systematic reviews from McMaster University (of over 40 people), widely known for its contributions to the field of research evidence synthesis. Their comments helped inform the emerging guidance.

The organising committee became the writing committee after the meeting. They developed the guidance arising from the meeting, developed the checklist and diagrams, added examples, and finalised the manuscript. The guidance was circulated to the larger group three times, with the PUGs panel providing extensive feedback. This feedback was all considered and carefully addressed by the writing committee. The writing committee provided the panel with the option of expressing any additional comments from the general or specific guidance in the report, and the option for registering their own view that might differ to the guidance formed and their view would be recorded in an annex. In the event, consensus was reached, and the annex was not required.

Definition of update

The PUGs panel defined an update of a systematic review as a new edition of a published systematic review with changes that can include new data, new methods, or new analyses to the previous edition. This expands on a previous definition of a systematic review update.[10] An update asks a similar question with regard to the participants, intervention, comparisons, and outcomes (PICO) and has similar objectives; thus it has similar inclusion criteria. These inclusion criteria can be modified in the light of developments within the topic area with new interventions, new standards, and new approaches. Updates will include a new search for potentially relevant studies and incorporate any eligible studies or data; and adjust the findings and conclusions as appropriate. Box 1 provides some examples.

Which systematic reviews should be updated and when?

Any group maintaining a portfolio of systematic reviews as part of their normative work, such as guidelines panels or Cochrane review groups, will need to prioritise which reviews to update. Box 2 presents the approaches used by the Agency for HealthCare Research and Quality (AHRQ) and Cochrane to prioritise which systematic reviews to update and when. Clearly, the responsibility for deciding which systematic reviews should be updated and when they will be updated will vary: it may be centrally organised and resourced, as with the AHRQ scientific resource centre (box 2). In Cochrane, the decision making process is decentralised to the Cochrane Review Group editorial team, with different approaches applied, often informally.

Box 1: Examples of what factors might change in an updated systematic review

- A systematic review of steroid treatment in tuberculosis meningitis used GRADE methods and split the composite outcome in the original review of death plus disability into its two components. This improved the clarity of the reviews findings in relation to the effects and the importance of the effects of steroids on death and on disability.[11]
- A systematic review of dihydroartemisinin-piperaquine (DHAP) for treating malaria was updated with much more detailed analysis of the adverse effect data from the existing trials as a result of questions raised by the European Medicines Agency. Because the original review included other comparisons, the update required extracting only the DHAP comparisons from the original review, and a modification of the title and the PICO.[12]
- A systematic review of atorvastatin was updated with simple uncontrolled studies.[13] This update allowed comparisons with trials and strengthened the review findings.[14]

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001520-00022

The PUGs panel recommended an individualised approach to updating, which used the procedures summarised in figure 1. The figure provides a status category, and some options for classifying reviews into each of these categories, and builds on a previous decision tool and earlier work developing an updating classification system.[15 16] We provide a narrative for each step.

Step 1: assess currency

Does the published review still address a current question?

An update is only worthwhile if the question is topical for decision making for practice, policy, or research priorities (fig 1). For agencies, people responsible for managing a portfolio of systematic reviews, there is a need to use both formal and informal horizon scanning. This type of scanning helps identify questions with currency, and can help identify those reviews that should be updated. The process could include monitoring policy debates around the review, media outlets, scientific (and professional) publications, and linking with guideline developers.

Has the review had good access or use?

Metrics for citations, article access and downloads, and sharing via social or traditional media can be used as proxy or indicators for currency and relevance of the review. Reviews that are widely cited and used could be important to update should the need arise. Comparable reviews that are never cited or rarely downloaded, for example, could indicate that they are not addressing a question that is valued, and might not be worth updating.

In most cases, updated reviews are most useful to stakeholders when there is new information or methods that result in a change in findings. However, there are some circumstances in which an up to date search for information is important for retaining the credibility of the review, regardless of whether the main findings would change or not. For example, key stakeholders would dismiss a review if a study is carried out in a relevant geographical setting but is not included; if a large, high profile study that might not change the findings is not included; or if an up to date search is required for a guideline to achieve credibility. Box 3 provides such examples. If the review does not answer a current question, the intervention has been superseded, then a decision can be made not to update and no further intelligence gathering is required (fig 1).

Did the review use valid methods and was it well conducted?

If the question is current and clearly defined, the systematic review needs to have used valid methods and be well conducted. If the review has vague inclusion criteria, poorly articulated outcomes, or inappropriate methods, then updating should not proceed. If the question is current, and the review has been cited or used, then it might be appropriate to simply start with a new protocol. The appraisal should take into account the methods in use when the review was done.

Step 2: identify relevant new methods, studies, and other information

Are there any new relevant methods?

If the question is current, but the review was done some years ago, the quality of the review might not meet current day standards. Methods have advanced quickly, and data extraction and understanding of the review process have become more sophisticated. For example:

- Methods for assessing risk of bias of randomised trials,[23] diagnostic test accuracy (QUADAS-2),[24] and observational studies (ROBINS-1).[25]
- Application of summary of findings, evidence profiles, and related GRADE methods has meant the characteristics of the intervention, characteristics of the participants, and risk of bias are more thoroughly and systematically documented.[26 27]
- Integration of other study designs containing evidence, such economic evaluation and qualitative research.[28]

There are other incremental improvements in a wide range of statistical and methodological areas, for example, in describing and taking into account cluster

Box 2: Examples of how different organisations decide on updating systematic reviews

Agency for Healthcare Research and Quality (US)

The AHRQ uses a needs based approach; updating systematic reviews depends on an assessment of several criteria:

1. Stakeholder impact
   - Interest from stakeholder partners (such as consumers, funders, guideline developers, clinical societies, James Lind Alliance)
   - Use and uptake (for example, frequency of citations and downloads)
   - Citation in scientific literature including clinical practice guidelines
2. Currency and need for update
   - New research is available
   - Review conclusions are probably dated
3. Update decision
   - Based on the above criteria, the decision is made to either update, archive, or continue surveillance.

Cochrane

Of over 50 Cochrane editorial teams, most but not all have some systems for updating, although this process can be informal and loosely applied. Most editorial teams draw on some or all of the following criteria:

1. Strategic importance
   - Is the topic a priority area (for example, in current debates or considered by guidelines groups)?
   - Is there important new information available?
2. Practicalities in organising the update that many groups take into account
   - Size of the task (size and quality of the review, and how many new studies or analyses are needed)
   - Availability and willingness of the author team
3. Impact of update
   - New research impact on findings and credibility
   - Consider whether new methods will improve review quality
4. Update decision
   - Priority to update, postpone update, class review as no longer requiring an update

ED_004886_00001520-00023

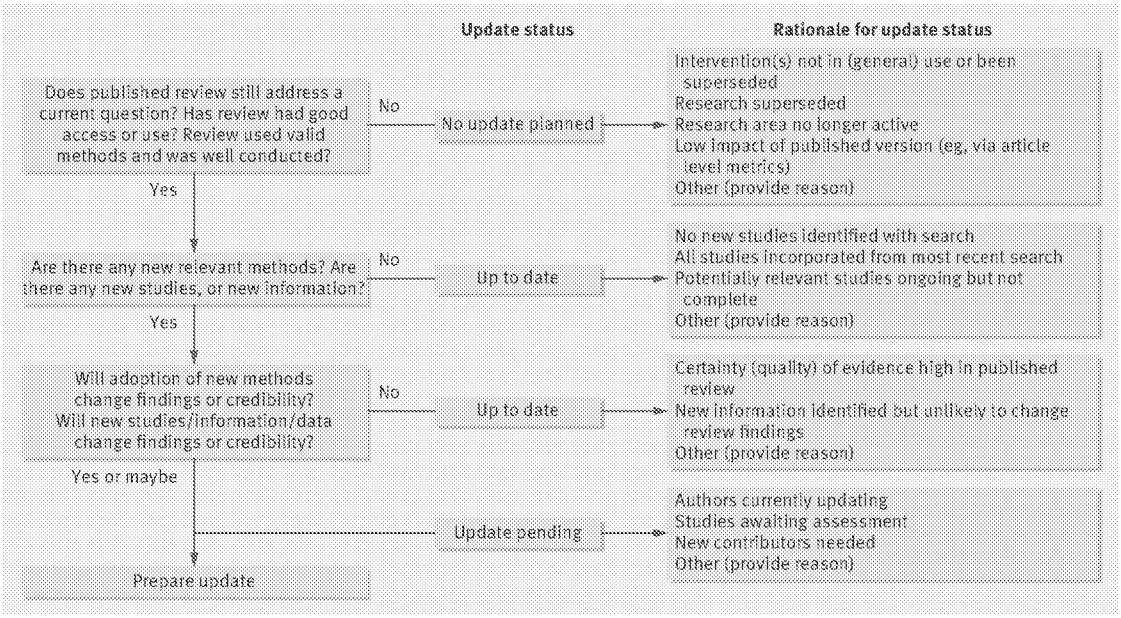


Fig 1 | Decision framework to assess systematic reviews for updating, with standard terms to report such decisions

randomised trials.[29] AMSTAR can assess the overall quality of a systematic review,[30] and the ROBIS tool can provide a more detailed assessment of the potential for bias.[31]

Are there any new studies or other information?

If an authoring or commissioning team wants to ensure that a particular review is up to date, there is a need for routine surveillance for new studies that are potentially relevant to the review, by searching and trial register inspection at regular intervals. This process has several approaches, including:

- Formal surveillance searching[32]
- Updating the full search strategies in the original review and running the searches
- Tracking studies in clinical trial and other registers
- Using literature appraisal services[33]
- Using a defined abbreviated search strategy for the update[34]
- Checking studies included in related systematic reviews.[35]

How often this surveillance is done, and which approaches to use, depend on the circumstances and the topic. Some topics move quickly, and the definition of "regular intervals" will vary according to the field and according to the state of evidence in the field. For example, early in the life of a new intervention, there might be a plethora of studies, and surveillance would be needed more frequently.

Box 3: Examples of a systematic review's currency

- The public is interested in vitamin C for preventing the common cold: the Cochrane review includes over 29 trials with either no or small effects, concluding good evidence of no important effects.[17] Assessment: still a current question for the public.
- Low osmolarity oral rehydration salt (ORS) solution versus standard solution for acute diarrhoea in children: the 2001 Cochrane review[18] led the World Health Organization to recommend ORS solution formula worldwide to follow the new ORS solution formula[19] and this has now been accepted globally. Assessment: no longer a current question.
- Routine prophylactic antibiotics with caesarean section: the Cochrane review reports clear evidence of maternal benefit from placebo controlled trials but no information on the effects on the baby.[20] Assessment: this is a current question.
- A systematic review published in the *Lancet* examined the effects of artemisinin based combination treatments compared with monotherapy for treating malaria and showed clear benefit.[21] Assessment: this established the treatment globally and is no longer a current question and no update is required.
- A Cochrane review of amalgam restorations for dental caries[22] is unlikely to be updated because the use of dental amalgam is declining, and the question is not seen as being important by many dental specialists. Assessment: no longer a current question.

Step 3: assess the effect of updating the review

Will the adoption of new methods change the findings or credibility?

Editors, referees, or experts in the topic area or methodologists can provide an informed view of whether a review can be substantially improved by application of current methodological expectations and new methods (fig 1). For example, a Cochrane review of iron supplementation in malaria concluded that there was "no significant difference between iron and placebo detected."[36] An update of the review included a GRADE assessment of the certainty of the evidence, and was able to conclude with a high degree of certainty that iron does not cause an excess of clinical malaria because the upper relative risk confidence intervals of harm was 1.0 with high certainty of evidence.[37]

Will the new studies, information, or data change the findings or credibility?

The assessment of new data contained in new studies and how these data might change the review is often used to determine whether an update should go ahead,

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001520-00024

Table 1 | Formal prediction tools: how potentially relevant new studies can affect review conclusions

Method	Description of approach	How it could be used	Advantages	Limitations	Validation
GRADE approach[38]	Considers whether the evidence certainty might change in the update (for example, because of lack of high certainty evidence, or because new evidence contradicts existing high certainty evidence). High certainty of evidence for critical outcomes could lower the priority for updating. Uncertainty in the review findings increases the need to include new studies[38]	Provides a benchmark by outcome to assess whether a new trial will improve the certainty of the evidence	Pragmatic. Many reviews already include GRADE	Requires GRADE to have been used in existing review or to complete an assessment according to GRADE	GRADE summary of findings tables or evidence profiles widely validated. Use of GRADE approach to prioritising updates requires further validation
Ottawa method[39-41]	A simple PubMed search (using the three largest and three most recent trials from the original review) to identify new studies. If new studies are found, then the method uses quantitative signals (eg, change in significance, effect size) to assess the likelihood that the new studies will change the review conclusion, thus triggering an update	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions; reasonable agreement with RAND method[39 42-44]
RAND method[45]	An abbreviated search of five major journals to identify new studies, and a search of the US Food and Drug Administration website and external expert judgment to determine the currency of the report findings	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions, and compares well with the Ottawa method[39 43 44]
Statistical prediction tool[15]	A multicomponent decision tool to assess whether there might be any new studies for the update. If new studies are identified, a statistical prediction tool estimates the probability that this will change the review conclusion	Ranks multiple systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires commercial software	Approach validated internally[46]; requires further external validation
Value of information analysis[47 48]	Builds on the statistical prediction tool approach[15] comparing the expected health gain from new evidence with its expected cost. The gain is calculated in terms of a reduction in expected loss from reduced uncertainty and the cost is measured in days required to update the review.	Ranks selected systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires some statistical knowledge	Approach validated internally; requires further external validation

Box 4: Examples of new information other than new trials being important

- The iconic Cochrane review of steroids in preterm labour was thought to provide evidence of benefit in infants, and this question no longer required new trials. However, a new large trial published in the *Lancet* in 2015 showed that in low and middle income countries, strategies to promote the uptake of neonatal steroids increased neonatal mortality and suspected maternal infection.[49] This information needs to somehow be incorporated into the review to maintain its credibility.
- A Cochrane review of community deworming in developing countries indicates that in recent studies, there is little or no effect.[50] The inclusion of a large trial of two million children confirmed that there was no effect on mortality. Although the incorporation of the trial in the review did not change the review's conclusions, the trial's absence would have affected the credibility of the review, so it was therefore updated.
- A new paper reporting long term follow-up data on anthracycline chemotherapy as part of cancer treatment was published. Although the effects from the outcomes remained essentially unchanged, apart from this longer follow-up, the paper also included information about the performance bias in the original trial, shifting the risk of bias for several outcomes from "unknown" to "high" in the Cochrane review.[51]

and the speed with which the update should be conducted. The appraisal of these new data can be carried out in different ways. Initially, methods focused on statistical approaches to predict an overturning of the current review findings in terms of the primary or desired outcome (table 1). Although this aspect is important, additional studies can add important information to a review, which is more than just changing the primary outcome to a more accurate and reliable estimate. Box 4 gives examples.

Reviews with a high level of certainty in the results (that is, when the GRADE assessment for the body of evidence is high) are less likely to change even with the addition of new studies, information, or data, by definition. GRADE can help guide priorities in whether to update, but it is still important to assess new studies that might meet the inclusion criteria. New studies can show unexpected effects (eg, attenuation of efficacy) or provide new information about the effects seen in different circumstances (eg, groups of patients or locations).

Other tools are specifically designed to help decision making in updating. For example, the Ottawa[39] and RAND[45] methods focus on identification of new evidence, the statistical predication tool[15] calculates the probability of new evidence changing the review conclusion, and the value of information analysis approach[52] calculates the expected health gain (table 1). As yet, there has been limited external validation of these tools to determine which approach would be most effective and when.

If potentially relevant studies are identified that have not previously been assessed for inclusion, authors or those managing the updating process need to assess whether including them might affect the conclusions of the review. They need to examine the weight and certainty of the new evidence to help determine whether an update is needed and how urgent that update is. The updating team can assess this informally by judging whether new studies or data are likely to substantively affect the review, for example, by altering the certainty in an existing comparison, or by generating new comparisons and analyses in the existing review.

New information can also include fresh follow-up data on existing included studies, or information on how the studies were carried out. These should be assessed in terms of whether they might change the review findings or improve its credibility (fig 1). Indeed, if any study has been retracted, it is important the authors assess the reasons for its retraction. In the case of data fabrication, the study needs to be removed

ED_004886_00001520-00025

from the analysis and this recorded. A decision needs to be made as to whether other studies by the same author should be removed from the review and other related reviews. An investigation should also be initiated following guidelines from the Committee on Publication Ethics (COPE). Additional published and unpublished data can become available from a wide range of sources—including study investigators, regulatory agencies and industry—and are important to consider.

Preparing for an update

Refresh background, objectives, inclusion criteria, and methods

Before including new studies in the review, authors need to revisit the background, objectives, inclusion criteria, and methods of the current review. In Cochrane, this is referred to as the protocol, and editors are part of this process. The update could range from simply endorsing the current question and inclusion criteria, through to full rewriting of the question, inclusion criteria and methods, and republishing the protocol. As a field progresses with larger and better quality trials rigorously testing the questions posed, it may be appropriate to exclude weaker study designs (such as quasi-randomised comparisons or very small trials) from the update (table 2). The PUGs panel recommended that a protocol refresh will require the authors to use the latest accepted methods of synthesis, even if this means repeating data extraction for all studies.

New authors and authorship

Updated systematic reviews are new publications with new citations. An authorship team publishing an update in a scientific or medical journal is likely to manage the new edition of a review in the same way as with any other publication, and follow the ICMJE authorship criteria.[56] If the previous author or author team steps down, then they should be acknowledged in the new version. However, some might perceive that their efforts in the first version warrant continued authorship, which may be valid. The management of authorship between versions can sometimes be complicated. At worst, it delays new authors completing an update and leads to long authorship lists of people from previous versions who probably do not meet ICMJE authorship criteria. One approach with updates including new authors is to have an opt-in policy for the existing authors: they can opt in to the new edition, provided that they make clear their contribution, and this is then agreed with the entire author team.

Although they are new publications, updates will generally include content from the published version. Changing licensing rights around systematic reviews to allow new authors of future updates to remix, tweak, or build on the contributions of the original authors of the published version (similar to the rights available via a Creative Commons licence; https://creativecommons.org) could be a more sustainable and simpler approach. This approach would allow systematic reviews to continue to evolve and build on the work of a range of authors over time, and for contributors to be given credit for contributions to this previous work.

Efficient searching

In performing an update, a search based on the search conducted for the original review is required. The updated search strategy will need to take into account changes in the review question or inclusion criteria, for example, and might be further adjusted based on knowledge of running the original search strategy. The search strategy for an update need not replicate the original search strategy, but could be refined, for example, based on an analysis of the yield of the original search. These new search approaches are currently undergoing formal empirical evaluation, but they may well provide much more efficient search strategies in the future. Some examples of these possible new methods for review updates are described in web appendix 2.

In reporting the search process for the update, investigators must ensure transparency for any previous versions and the current update, and use an adapted flow diagram based on PRISMA reporting (preferred reporting items for systematic reviews and meta-analyses).[57] The search processes and strategies for the update must be adequately reported such that they could be replicated.

Table 2 | Refresh background, objectives, inclusion criteria, and methods

Protocol section	Appraisal points
Background and research question	• Review and update background section, including supporting references to take account of any changes that may have occurred. This should include updating any new information and current policy debates on the topic. • Assess whether the current review question remains relevant to patients and practice.
Inclusion criteria	• Consider whether the existing PICO(s) remain(s) current, in the light of new knowledge. • Identify any new understanding of definition of patient populations. • Identify new interventions, or those that have been withdrawn, that are no longer in use. • Identify any changes in usual care standards. • Check for standardised core outcomes sets, such as those developed in collaboration with the core outcome measures in effectiveness trials (COMET) initiative (www.comet-initiative.org) or by guideline groups since the original review. • Check for any relevant patient reported outcomes to include subsequent to the original review. • Consider any new studies with less risk of bias that might warrant a stricter study design inclusion criteria (where the older version, when there was a dearth of evidence, included observational or quasi-randomised comparisons).
Methods	• Appraise and update the methods pending relevant methodological advancements or developments. For example, if (1) there are new tools for assessing the risk of bias of individual studies or appraising the quality of a body of evidence (eg, GRADE); or (2) new and efficient search approaches are feasible, such as a targeted approach to searching, taking into account the quality of the original search, and ensuring that the search for the update is of high quality. • Update or include a summary of findings table, which is recommended for all systematic reviews, because it improves the clarity, understanding, and interpretation of the findings of a systematic review, and rapidly reduces the amount of time readers require to find key information.[53-55]

ED_004886_00001520-00026

Peer review

Systematic reviews published for the first time in peer reviewed journals are by definition peer reviewed, but practice for updates remains variable, because an update might have few changes (such as an updated search but no new studies found and therefore included) or many changes (such as revise methods and inclusion of several new studies leading to revised conclusions). Therefore, and to use peer reviewers' time most effectively, editors need to consider when to peer review an update and the type of peer reviewer most useful for a particular update (for example, topic specialist, methodologist). The decision to use peer review, and the number and expertise of the peer reviewers could depend on the nature of the update and the extent of any changes to the systematic review as part of an editor assessment. A change in the date of the search only (where no new studies were identified) would not require peer review (except, arguably, peer review of the search), but the addition of studies that lead to a change in conclusions or significant changes to the methods would require peer review. The nature of the peer review could be described within the published article.

Reporting changes

Authors should provide a clear description of the changes in approach or methods between different editions of a review. Also, authors need to report the differences in findings between the original and updated edition to help users decide how to use the new edition. The approach or format used to present the differences in findings might vary with the target user group.[58] Publishers need to ensure that all previous versions of the review remain publically accessible.

Updates can range from small adjustments to reviews being completely rewritten, and the PUGs panel spent some time debating whether the term "new edition" would be a better description than "update." However, the word "update" is now in common parlance and changing the term, the panel judged, could cause confusion. However, the debate does illustrate that an update could represent a review that asks a similar question but has been completely revised.

Technology and innovation

The updating of systematic review is generally done manually and is time consuming. There are opportunities to make better use of technology to streamline the updating process and improve efficiency (table 3). Some of these tools already exist and are in development or in early use, and some are commercially available or freely available. The AHRQ's evidence based practice centre team has recently published tools for searching and screening, and will provide an assessment of the use, reliability, and availability of these tools.[63]

Other developments, such as targeted updates that are performed rapidly and focus on updating only key components of a review, could provide different approaches to updating in the future and are being piloted and evaluated.[64] With implementation of these various innovations, the longer term goal is for "living" systematic reviews, which identify and incorporate information rapidly as it evolves over time.[60]

Concluding remarks

Updating systematic reviews, rather than addressing the same question with a fresh protocol, is generally

Table 3 | Technological innovations to improve the efficiency of updating systematic reviews

Innovation	Description	Application	Examples of software and projects,* and current status
Integrated software	Integration of applying inclusion criteria, review management systems, statistical packages, and GRADE	To facilitate greater efficiencies in review production, including their updates	Covidence (www.covidence.org): free/pay† EPPI reviewer (http://eppi.ioe.ac.uk/cms/Default.aspx): pay DistillerSR (https://distillercer.com/products/distillersr-systematic-review-software/): pay Cochrane Review Manager (http://tech.cochrane.org/revman): free/pay† GRADEpro GDT (http://gradepro.org/): free Rayyan (http://rayyan.qcri.org/): free
Systematic review data repositories	Repositories store information from review (eg, data abstraction forms and the evidence tables)	Improve updating efficiency for new or existing teams as the data abstraction forms, evidence tables, and populating data from the original review are available	Agency of Health Care Research and Quality systematic review data repository (http://srdr.ahrq.gov): operational GRADE database of evidence profiles and evidence to decision frameworks (http://dbep.gradepro.org/): operational
Semi-automation	Machine learning techniques to use alongside human efforts	Finding studies and extracting data could benefit from semi-automation creating time efficiencies[59]	RobotReviewer (http://vortext.systems/robotreviewer): free Cochrane project transform—identification of studies (http://community.cochrane.org/tools/project-coordination-and-support/transform)
Crowdsourcing	Use of volunteers to assist systematic review authors with discrete tasks	Individuals from the "crowd" assist with tasks (identifying and screening studies, translating articles, data extraction) to help in new review production and updates[60-62]	Cochrane Project Transform—crowdsourcing (link as above)
Publication linkage	Ability to link trial registration, trial publications, and reviews citing them will help transparency	This initiative could help identify studies for systematic reviews and could also show the relation between systematic review updates	A cross publisher initiative, CrossRef, is coordinating a threaded publications/linked clinical trial reports initiative to link a clinical trial report (with a trial registration number) report and derivative publications, including reviews (www.crossref.org): operational
Data linkage	Increase links between data, existing software, and reviews	To improve identification and reuse of data for review production and dissemination	Cochrane (http://linkeddata.cochrane.org/): proof of principle example at production phase, but mostly linkage projects at exploratory phase

*Further information can be located on the SR Toolbox site (http://systematicreviewtools.com/).
†Free to Cochrane contributors; other users pay.

ED_004886_00001520-00027

more efficient and allows incremental improvement over time. Mechanical rules appear unworkable, but there is no clear unified approach on when to update, and how implement this. This PUGs panel of authors, editors, statisticians, information specialists, other methodologists, and guideline developers brought together current thinking and experience in this area to provide guidance.

Decisions about whether and when to update a systematic review are judgments made at a point in time. They depend on the currency of the question asked, the need for updating to maintain credibility, the availability of new evidence, and whether new research or new methods will affect the findings.

Whether the review uses current methodological standards is important in deciding if the update will influence the review findings, quality, reliability, or credibility sufficiently to justify the effort in updating it. Those updating systematic reviews to author clinical practice guidelines might consider the influence of new study results in potentially overturning the conclusions of an existing review. Yet, even in cases where new study findings do not change the primary outcome measure, new studies can carry important information about subgroup effects, duration of treatment effects, and other relevant clinical information, enhancing the currency and breadth of review results.

An update requires appraisal and revision of the background, question, inclusion criteria, and methods of the existing review and the existing certainty in the evidence. In particular, methods might need to be updated, and search strategies reconsidered. Authors of updates need to consider inputs to the current edition, and follow ICMJE criteria regarding authorship.[56]

The PUGs panel proposed a decision framework (fig 1), with terms and categories for reporting the decisions made for updating procedures for adoption by Cochrane and other stakeholders. This framework includes journals publishing systematic review updates and independent authors considering updates of existing published reviews. The panel developed a checklist to help judgements about when and how to update.

The current emphasis of authors, guideline developers, Cochrane, and consequently this guidance has been on effects reviews. The checklists and guidance here still applies to other types of systematic reviews, such as those on diagnostic test accuracy, and this guidance will need adapting. Accumulative experience and methods development in reviews other than those of effects are likely to help refine guidance in the future.

This guidance could help groups identify and prioritise reviews for updating and hence use their finite resources to greatest effect. Software innovation and new management systems are being developed and in early use to help streamline review updates in the coming years.

AUTHOR AFFILIATIONS

[1]Cochrane Infectious Diseases Group, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK
[2]Oxford Clinical Trials Research Unit, University of Oxford, Oxford, UK
[3]Cochrane Editorial Unit, Cochrane Central Executive, London, UK
[4]Department of Clinical Epidemiology and Biostatistics and Department of Medicine, McMaster University, Hamilton, ON, Canada
[5]Cochrane GRADEing Methods Group, Ottawa, ON, Canada
[6]Department of Internal Medicine, American University of Beirut, Beirut, Lebanon
[7]Department of Mathematics and Statistics, McMaster University
[8]Evidence-based Practice Center Program, Agency for Healthcare and Research Quality, Rockville, MD, USA
[9]Centre for Reviews and Dissemination, University of York, York, UK
[10]Cochrane Upper Gastrointestinal and Pancreatic Diseases Group, Hamilton, ON, Canada
[11]Lefebvre Associates, Oxford, UK
[12]Kaiser Permanente National Guideline Program, Portland, OR, USA
[13]Iberoamerican Cochrane Centre, Barcelona, Spain
[14]Cochrane Informatics and Knowledge Management, Cochrane Central Executive, Freiburg, Germany
[15]Plymouth University Peninsula School of Dentistry, Plymouth, UK
[16]Department of Clinical Policy, American College of Physicians,Philadelphia, PA, USA
[17]Guidelines International Network, Pitlochry, UK
[18]Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
[19]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[20]Biostatistics Unit, Centre for Evaluation, McMaster University, Hamilton, ON, Canada
[21]Centre for Statistics in Medicine, University of Oxford, Oxford, UK
[22]University of Ottawa, Ottawa, ON, Canada
[23]Cochrane Airways Group, Population Health Research Institute, St George's, University of London, London, UK
[24]Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, UK

Contributors: HJS initiated the workshop. JC, SH, PG, HM, and HJS organised the materials and the agenda. SH wrote up the proceedings. PG wrote the paper from the proceedings and coordinated the development of the final guidance; JC, SH, HM, and HJS were active in the finalising of the guidance. All PUGs authors contributed to three rounds of manuscript revision.

Funding: Attendance at this meeting, for those attendees not directly employed by Cochrane, was not funded by Cochrane beyond the reimbursement of out of pocket expenses for those attendees for whom this was appropriate. Expenses were not reimbursed for US federal government attendees, in line with US government policy. Statements in the manuscript should not be construed as endorsement by the US Agency for Healthcare Research and Quality or the US Department of Health and Human Services.

Competing interests: All participants have a direct or indirect interest in systematic reviews and updating as part of their job or academic career. Most participants contribute to Cochrane, whose mission includes a commitment to the updating of its systematic review portfolio. JC, HM, RM, CM, KS-W, and MT are, or were at that time, employed by the Cochrane Central Executive.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 3.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/3.0/.

1 Shekelle PG, Ortiz E, Rhodes S, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated?*JAMA* 2001;286:1461-7. doi:10.1001/jama.286.12.1461.
2 Claxton K, Cohen JT, Neumann PJ. When is evidence sufficient?*Health Aff (Millwood)* 2005;24:93-101. doi:10.1377/hlthaff.24.1.93.
3 Fenwick E, Claxton K, Sculpher M, et al. *Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Paper 179.*Centre for Health Economics, University of York, 2000.
4 Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty—when is there sufficient evidence?*Value Health* 2005;8:433-46. doi:10.1111/j.1524-4733.2005.00033.x.

ED_004886_00001520-00028

5 Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *J Health Serv Res Policy* 1997;2:26-30.
6 Wilson E, Abrams K. From evidence based economics to economics based evidence: using systematic review to inform the design of future research. In: Shemilt I, Mugford M, Vale L, et al. eds. *Evidence based economics*.Blackwell Publishing, 2010doi:10.1002/9781444320398.ch12.
7 Chalmers I, Enkin M, Keirse MJ. Preparing and updating systematic reviews of randomized controlled trials of health care. *Milbank Q* 1993;71:411-37. doi:10.2307/3350409.
8 Higgins J, Green S, Scholten R. Chapter 3. *Maintaining reviews: updates, amendments and feedback: Version 5.1.0 (updated March 2011)*.Cochrane Collaboration, 2011.
9 Cochrane. Editorial and publishing policy resource. http://community.cochrane.org/editorial-and-publishing-policy-resource. 2016.
10 Moher D, Tsertsvadze A. Systematic reviews: when is an update an update?*Lancet* 2006;367:881-3. doi:10.1016/S0140-6736(06)68358-X.
11 Prasad K, Singh MB, Ryan H. Corticosteroids for managing tuberculous meningitis. *Cochrane Database Syst Rev* 2016;4:CD002244.
12 Zani B, Gathu M, Donegan S, Olliaro PL, Sinclair D. Dihydroartemisinin-piperaquine for treating uncomplicated *Plasmodium falciparum* malaria. *Cochrane Database Syst Rev* 2014;1:CD010927.
13 Adams SP, Tsang M, Wright JM. Lipid lowering efficacy of atorvastatin. *Cochrane Database Syst Rev* 2012;12:CD008226.
14 Higgins J. Convincing evidence from controlled and uncontrolled studies on the lipid-lowering effect of a statin. *Cochrane Database Syst Rev* 2012;12:ED000049.
15 Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ* 2013;347:f7191. doi:10.1136/bmj.f7191.
16 MacLehose H, Hilton J, Tovey D, et al. The Cochrane Library: revolution or evolution? Shaping the future of Cochrane content. Background paper for The Cochrane Collaboration's Strategic Session Paris, France, 18 April 2012. http://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/2012-CC-strategic-session_full-report.pdf.
17 Hemilä H, Chalker E. Vitamin C for preventing and treating the common cold. *Cochrane Database Syst Rev* 2013;(1):CD000980.
18 Hahn S, Kim S, Garner P. Reduced osmolarity oral rehydration solution for treating dehydration caused by acute diarrhoea in children. *Cochrane Database Syst Rev* 2002;(1):CD002847.
19 World Health Organization (WHO). Reduced osmolarity oral rehydration salts (ORS) formulation. A report from a meeting of Experts jointly organized by UNICEF and WHO. New York: Child and Adolescent Health and Development, 18 July 2001 http://apps.who.int/iris/bitstream/10665/67322/1/WHO_FCH_CAH_01.22.pdf.
20 Smaill FM, Grivell RM. Antibiotic prophylaxis versus no prophylaxis for preventing infection after cesarean section. *Cochrane Database Syst Rev* 2014;(10):CD007482.
21 Adjuik M, Babiker A, Garner P, Olliaro P, Taylor W, White N. International Artemisinin Study Group. Artesunate combinations for treatment of malaria: meta-analysis. *Lancet* 2004;363:9-17. doi:10.1016/S0140-6736(03)15162-8.
22 Agnihotry A, Fedorowicz Z, Nasser M. Adhesively bonded versus non-bonded amalgam restorations for dental caries. *Cochrane Database Syst Rev* 2016;3:CD007517.
23 Higgins JP, Altman DG, Gøtzsche PC, et al. Cochrane Bias Methods Group Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928.
24 Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36. doi:10.7326/0003-4819-155-8-201110180-00009.
25 Sterne JAC, Higgins JPT, Reeves BC; on behalf of the development group for ROBINS-I. A tool for assessing risk of bias in non-randomized studies of interventions, version 7. March 2016. www.riskofbias.info.
26 Guyatt GH, Oxman AD, Vist GE, et al. GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD.
27 Schünemann HJ. Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision?*J Clin Epidemiol* 2016;75:6-15. doi:10.1016/j.jclinepi.2016.03.018.
28 Gough D. Qualitative and mixed methods in systematic reviews. *Syst Rev* 2015;4:181. doi:10.1186/s13643-015-0151-y.
29 Richardson M, Garner P, Donegan S. Cluster randomised trials in Cochrane reviews: evaluation of methodological and reporting practice. *PLoS One* 2016;11:e0151818. doi:10.1371/journal.pone.0151818.
30 Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10. doi:10.1186/1471-2288-7-10.
31 Whiting P, Savović J, Higgins JP, et al. ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005.
32 Sampson M, Shojania KG, McGowan J, et al. Surveillance search techniques identified the need to update systematic reviews. *J Clin Epidemiol* 2008;61:755-62. doi:10.1016/j.jclinepi.2007.10.003.
33 Hemens BJ, Haynes RB. McMaster Premium LiteratUre Service (PLUS) performed well for identifying new studies for updated Cochrane reviews. *J Clin Epidemiol* 2012;65:62-72.e1. doi:10.1016/j.jclinepi.2011.02.010.
34 Sagliocca L, De Masi S, Ferrigno L, Mele A, Traversa G. A pragmatic strategy for the review of clinical evidence. *J Eval Clin Pract* 2013;19:689-96. doi:10.1111/jep.12020.
35 Rada G, Peña J, Capurro D, et al. How to create a matrix of evidence in Epistemonikos. Abstracts of the 22nd Cochrane Colloquium; Evidence-informed public health: opportunities and challenges; Hyderabad, India. *Cochrane Database Syst Rev* 2014;suppl 1:132.
36 Okebe JU, Yahav D, Shbita R, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2011;(10):CD006589.
37 Neuberger A, Okebe J, Yahav D, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2016;2:CD006589.
38 Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401-6. doi:10.1016/j.jclinepi.2010.07.015.
39 Chung M, Newberry SJ, Ansari MT, et al. Two methods provide similar signals for the need to update systematic reviews. *J Clin Epidemiol* 2012;65:660-8. doi:10.1016/j.jclinepi.2011.12.004.
40 Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med* 2007;147:224-33. doi:10.7326/0003-4819-147-4-200708210-00179.
41 Shojania K, Sampson M, Ansari M, et al. *Updating systematic reviews; AHRQ technical reviews; report no 07-0087*.Agency for Healthcare Research and Quality, 2007.
42 Pattanittum P, Laopaiboon M, Moher D, Lumbiganon P, Ngamjarus C. A comparison of statistical methods for identifying out-of-date systematic reviews. *PLoS One* 2012;7:e48894. doi:10.1371/journal.pone.0048894.
43 Shekelle PG, Motala A, Johnsen B, Newberry SJ. Assessment of a method to detect signals for updating systematic reviews. *Syst Rev* 2014;3:13. doi:10.1186/2046-4053-3-13.
44 Shekelle PG, Newberry SJ, Wu H, et al. *Identifying signals for updating systematic reviews: a comparison of two methods; report no 11-EHC042-EF*.Agency for Healthcare Research and Quality, 2011.
45 Shekelle P, Newberry S, Maglione M, et al. *Assessment of the need to update comparative effectiveness reviews: report of an initial rapid program assessment (2005-2009)*.Agency for Healthcare Research and Quality, 2009.
46 Tovey D, Marshall R, Bazian, Hopewell S, Rader T. Fit for purpose: centralised updating support for high-priority Cochrane Reviews; National Institute for Health Research Cochrane-National Health Service Engagement Award Scheme, July 2011. https://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/10_4000_01%20Fit%20for%20purpose%20-%20centralised%20updating%20support%20for%20high%20priority%20Cochrane%20Reviews%20FINAL%20REPORT.pdf.
47 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341-64. doi:10.1016/S0167-6296(98)00039-3.
48 Wilson EC. A practical guide to value of information analysis. *Pharmacoeconomics* 2015;33:105-21. doi:10.1007/s40273-014-0219-x.
49 Althabe F, Belizán JM, McClure EM, et al. A population-based, multifaceted strategy to implement antenatal corticosteroid treatment versus standard care for the reduction of neonatal mortality due to preterm birth in low-income and middle-income countries: the ACT cluster-randomised trial. *Lancet* 2015;385:629-39. doi:10.1016/S0140-6736(14)61651-2.
50 Taylor-Robinson D, Maayan N, Soares-Weiser K, et al. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin and school performance. *Cochrane Database Syst Rev* 2012;(11):CD000371.
51 van Dalen EC, van der Pal HJ, Kremer LC. Different dosage schedules for reducing cardiotoxicity in people with cancer receiving anthracycline chemotherapy. *Cochrane Database Syst Rev* 2016;3:CD005008.
52 Wilson E. on behalf of the Cochrane Priority Setting and Campbell & Cochrane Economics Methods Groups. Which study when? Proof of concept of a proposed automated tool to help decision which reviews to update first. *Cochrane Database Syst Rev* 2014;suppl 2:29-31.
53 Rosenbaum SE, Glenton C, Nylund HK, Oxman AD. User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews. *J Clin Epidemiol* 2010;63:607-19. doi:10.1016/j.jclinepi.2009.12.013.
54 Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. *J Clin Epidemiol* 2010;63:620-6. doi:10.1016/j.jclinepi.2009.12.014.

ED_004886_00001520-00029

55 Vandvik PO, Santesso N, Akl EA, et al. Formatting modifications in GRADE evidence profiles improved guideline panelists comprehension and accessibility to information. A randomized trial. *J Clin Epidemiol* 2012;65:748-55. doi:10.1016/j.jclinepi.2011.11.013.
56 International Committee of Medical Journal Editors (ICMJE). Defining the role of authors and contributors. 2016. www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.
57 Stovold E, Beecher D, Foxlee R, Noel-Storr A. Study flow diagrams in Cochrane systematic review updates: an adapted PRISMA flow diagram. *Syst Rev* 2014;3:54. doi:10.1186/2046-4053-3-54.
58 Newberry SJ, Shekelle PG, Vaiana M, et al. *Reporting the findings of updated systematic reviews of comparative effectiveness: how do users want to view new information? report no 13-EHC093-EF* Agency for Healthcare Research and Quality, 2013.
59 Marshall IJ, Kuiper J, Wallace BC. Automating risk of bias assessment for clinical trials BCB'14. Proceedings of the 5th ACM conference on Bioinformatics, computational biology, and health informatics. 2014:88-95. http://thirdworld.nl/automating-risk-of-bias-assessment-for-clinical-trials.
60 Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014;11:e1001603. doi:10.1371/journal.pmed.1001603.
61 Elliott J, Sim I, Thomas J, et al. #CochraneTech: technology and the future of systematic reviews. *Cochrane Database Syst Rev* 2014;(9):ED000091.
62 Cochrane. Project transform: the Cochrane Collaboration. 2016. http://community.cochrane.org/tools/project-coordination-and-support/transform.
63 Paynter R, Bañez L, Berlinerm E, et al. EPC methods: an exploration of the use of text-mining software in systematic reviews. Research white paper. AHRQ publication 16-EHC023-EF. Agency for Healthcare Research and Quality, April 2016. https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2214.
64 Soares-Weiser K, Marshall R, Bergman H, et al. Updating Cochrane Reviews: results of the first pilot of a focused update. *Cochrane Database Syst Rev* 2014;suppl 1:31-3.

© BMJ Publishing Group Ltd 2016

Web appendix: Supplementary material

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe

September 19, 2017

Comments from Academics, Scientists and Clinicians on the Risk Evaluation Scoping Efforts Under TSCA for Ten Chemical Substances

Submitted online via *Regulations.gov* to dockets: EPA-HQ-OPPT-2016-0736, EPA-HQ-OPPT-2016-0741, EPA-HQ-OPPT-2016-0733, EPA-HQ-OPPT-2016-0723, EPA-HQ-OPPT-2016-0735, EPA-HQ-OPPT-2016-0742, EPA-HQ-OPPT-2016-0743, EPA-HQ-OPPT-2016-0732, EPA-HQ-OPPT-2016-0725, EPA-HQ-OPPT-2016-0737

These comments are submitted on behalf of the undersigned academic, scientists, and clinicians. We declare collectively that we have no direct or indirect financial or fiduciary interest in any chemical under consideration in these risk evaluations. The co-signers' institutional affiliations are included for identification purposes only and do not necessarily imply any institutional endorsement or support, unless indicated otherwise.

We appreciate the opportunity to provide written comments on the scope of risk evaluations for the first ten chemical substances for risk evaluations pursuant to the Toxic Substances Control Act (TSCA), as amended by the Frank R. Lautenberg Chemical Safety of the 21st Century Act (Lautenberg TSCA). Collectively, these chemicals represent an aggregate production volume of more than 1 billion pounds a year in 2015.[1] Some of these chemicals have assessments, and in some cases even restrictions, under other federal programs – but none of these other programs has the mandate given to EPA under the new TSCA: to comprehensively evaluate chemicals and ensure that they do not pose an unreasonable risk to human health and the environment, with special consideration to those most vulnerable amongst us. Therefore, the task ahead for EPA is critical.

These first ten evaluations are also consequential because they will be precedent setting for the implementation of evaluation of science under TSCA. The consequent health impacts of EPA's decisions – for better or worse – will be borne by generations of American children, workers, families, and communities. With so much at stake, we welcome EPA's engagement with the public in this process and we offer EPA concrete approaches to embed the most current scientific principles in its methods to assess the hazards and risks of environmental chemicals.

Our comments address the following main points:

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**
2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**
3. **EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations**

[1] This is the aggregate production volume estimate for the 8 chemicals with production volume information available. For asbestos and pigment violet 29, manufacturers/ importers claimed production volumes as confidential business information (CBI).

ED_004886_00001537-00001

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**
5. **For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.**
6. **Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.**

We are appreciative of the opportunity to provide public input and we look forward to continuing to participate in such opportunities in the near future. Please do not hesitate to contact us with any questions regarding these comments.

Sincerely,

Veena Singla, PhD; Juleen Lam, PhD; and Tracey Woodruff, PhD
Program on Reproductive Health and the Environment
University of California, San Francisco

Patricia D. Koman, MPP, PhD
President and Senior Health Scientist
Green Barn Research

Kathy Attar, MPH
Toxics Program Manager
Physicians for Social Responsibility

Sheila Brear, BDS
Associate Dean of Academic Affairs, School of Dentistry
University of California, San Francisco

Phil Brown, PhD
University Distinguished Professor of Sociology and Health Sciences
Northeastern University

Adelita G. Cantu, PhD, RN
Associate Professor
Alliance of Nurses for Healthy Environments

Courtney Carignan, PhD
Assistant Professor
Department of Food Science and Human Nutrition and Department of Pharmacology and Toxicology
Michigan State University

ED_004886_00001537-00002

Jeanne A. Conry, MD, PhD
President, Environmental Health Leadership Foundation
Past President, The American Congress of Obstetricians and Gynecologists

Carl F. Cranor, PhD
Distinguished Professor of Philosophy
Faculty Member, Environmental Toxicology Graduate Program
Department of Philosophy
University of California, Riverside

Adam M. Finkel, ScD, CIH
Clinical Professor of Environmental Health Sciences, University of Michigan School of Public Health
Executive Director, Penn Program on Regulation, University of Pennsylvania Law School

Susan J. Fisher, PhD
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Robert M. Gould, MD
Associate Adjunct Professor
Program on Reproductive Health and the Environment
University of California, San Francisco
Past President, Physicians for Social Responsibility

Alycia Halladay, PhD
Chief Science Officer
Autism Science Foundation

Jean-Marie Kauth, PhD
Associate Professor
Benedictine University

Detlef R.U. Knappe, PhD
Professor
Department of Civil, Construction, and Environmental Engineering
North Carolina State University

Erica Koustas, PhD
Scientific Consultant
Program on Reproductive Health and the Environment
University of California, San Francisco

Hal C. Lawrence, III, MD, FACOG*
Executive Vice President and CEO
American Congress of Obstetricians and Gynecologists

ED_004886_00001537-00003

Heather Patisaul, PhD
Professor, Biological Sciences
Center for Human Health and the Environment, WM Keck Center for Behavioral Biology
North Carolina State University

Joshua F. Robinson, PhD
Assistant Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

Ted Schettler MD, MPH
Science Director
Science and Environmental Health Network

Patrice Sutton, MPH
Research Scientist
Program on Reproductive Health and the Environment
University of California, San Francisco

Aolin Wang, PhD
Postdoctoral Scholar
Program on Reproductive Health and the Environment
University of California, San Francisco

Nsedu Obot Witherspoon, MPH
Executive Director
Children's Environmental Health Network

Lauren Zajac, MD, MPH
Assistant Professor
Department of Environmental Medicine and Public Health
Icahn School of Medicine at Mount Sinai

Marya G. Zlatnik, MD, MMS
Professor
Department of Obstetrics, Gynecology and Reproductive Sciences
University of California, San Francisco

*indicates organizational support

ED_004886_00001537-00004

DETAILED COMMENTS

1. **EPA should improve its literature search and systematic review strategies to strengthen its evaluations and increase transparency.**

Overall, we strongly commend the EPA for its efforts to utilize a systematic and transparent method of research synthesis to reach a concise, strength of evidence conclusion about the human health hazard resulting from exposures to these ten chemicals. Efforts to integrate systematic review methods, including the explicit development of search terms, strategies, and inclusion/exclusion criteria beforehand, is relatively new in EPA's chemical assessment and as such, we applaud the EPA for this and its general improvements in its hazard assessment methodology. These scoping documents generally provide an important infrastructure for outlining EPA's screening approach for identifying relevant references and to document decisions made in the process of identifying the body of scientific literature that will be evaluated in the chemical assessments.

To improve on this document and advance EPA's uptake of systematic review methods of research synthesis, we identify the following opportunities for improvement:

EPA should not exclude studies based on language. EPA's search strategy is limited to English-only studies. The exclusive reliance on English-language studies may not represent the entire body of available evidence, and studies have suggested that language bias might lead to erroneous conclusions.[2] Furthermore, when considering the inclusion or update of an existing systematic review, studies have found that language-inclusive systematic reviews (including studies in languages other than English) were of the highest quality, compared with other types of reviews.[3] Online translation tools are readily available to allow screeners to quickly evaluate study abstracts for relevance, and therefore we recommend EPA to incorporate non-English language studies in their screening and not simply exclude these potentially relevant papers.

EPA should provide exclusion reasons for off topic citations. In the Bibliography Supplemental File for the Scope Documents, EPA has provided lists of bibliographic citations that were identified and screened from the initial literature search and the initial categorization of whether citations were *on topic* or *off topic.* We recommend EPA additionally provide exclusion reasons that were used to come to the conclusion that each citation was *off topic*, as this is a standard recommendation to fulfill transparency in documenting and reporting all decisions made in the study selection process.[4] This is particularly important as EPA has proposed to do its screening in Distiller, proprietary software that presumably will

---

[2] Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., Mierzwinski-Urban, M., Clifford, T., Hutton, B. and Rabb, D., 2012. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *International journal of technology assessment in health care*, *28*(2), pp.138-144.

[3] Moher, D., Pham, B., Lawson, M.L. and Klassen, T.P., 2003. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess*, *7*(41), pp.1-90.

[4] McDonagh M et a. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Agency for Healthcare Research and Quality. 2008.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J. and Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS medicine*, *6*(7), p.e1000100.

Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. March 2011. Available from: http://handbook-5-1.cochrane.org/chapter_7/7_2_5_selecting_excluded_studies.htm

ED_004886_00001537-00005

not be made publicly available, raising concerns regarding the transparency and reproducibility of this screening step.

EPA should consider other tools for systematic review. EPA has also proposed to extract data results in the DRAGON software. We strongly encourage EPA to also consider other potential software tools that have been developed and actively incorporated into the process of systematic review, such as Swift Reviewer,[5] Active Screener,[6] HAWC (Health Assessment Workplace Collaborative).[7] These tools will help to ensure consistent and transparent execution and presentation of reviews and increase transparency of EPA assessment. Furthermore, we urge EPA to work with the National Toxicology Program and other organizations involved in these efforts in an ongoing basis to develop these and other open source tools to train scientists in their use. We believe that such infrastructure development will be critical to increasing the efficiency of chemical assessments and to expedite uptake of systematic reviews in environmental health.

EPA should have two independent reviewers for screening steps. EPA has outlined its process for screening title and abstracts of papers as having a single reviewer reviewing papers to determine whether the study is *on-topic* or *off-topic*. As part of this process, a senior-level technical expert in the topic area of interest reviewed the appropriateness of the assigned tags for "the first batch of studies" and provided feedback to the screener. Senior-level technicians also provided feedback and guidance on specific references to the individual screeners as needed during the screening and tagging process. From the description of this process, it does not appear that two independent reviewers screen all titles and abstracts for potential inclusion. Using two independent reviewers is a standard approach in systematic reviews and therefore we strongly recommend that EPA include a second independent reviewer within this process to ensure that **all** studies are screened by two reviewers at each step (title and abstract and full text).

EPA should clearly document decisions related to the identification and search. For example, it was unclear how many studies were included in the first batch of studies reviewed by the senior-level technician—these decisions should be clearly specified beforehand as to the number (or percent) that will be reviewed by this independent reviewer. Furthermore, it is unclear how many studies the senior-level technical experts are reviewing generally as to their additional feedback and guidance to individual screeners. This should be more clearly stated and described beforehand in these protocols.
We recommend EPA broaden the set of studies that are initially screened in the first batch to ensure consistency across reviewers and demonstrated understanding of protocol instructions by all reviewers before moving on to screening the remaining records. It is stated in the Gray Literature Search Results that individual screeners would screen and tag 10 references that would be then independently reviewed by the senior-level technical expert. However, this does not seem to be an adequate number of studies as it is a small number relative to the expected number of records that will ultimately be screened.

EPA should clarify how it will handle discrepancies in the inclusion/exclusion and tagging process. As it is stated in its current protocol, it appears that the senior-level technical expert has the final say in

---

[5] Howard, B.E., Phillips, J., Miller, K., Tandon, A., Mav, D., Shah, M.R., Holmgren, S., Pelch, K.E., Walker, V., Rooney, A.A. and Macleod, M., 2016. SWIFT-Review: a text-mining workbench for systematic review. *Systematic reviews*, *5*(1), p.87.
[6] https://www.sciome.com/swift-activescreener/
[7] https://hawcproject.org/about/

determining the final inclusion/exclusion decision and tagging, for the subset of studies they evaluate. However, this should be clarified and we also highly recommend that a third party reviewer be incorporated as an arbiter for these decisions if consensus cannot be reached between the two reviewers, as is typically standard in systematic reviews.

EPA should stratify its exclusion criteria separately at the title and abstract and full-text screening steps. It is likely that title and abstracts of papers would not contain sufficient detail to evaluate all exclusion criteria—many of these would likely only be identified in the full-text of the paper. To increase the efficiency of the screening process, it would help to create a subset of exclusion criteria most relevant when screening the title and abstracts of records versus the larger set of exclusion criteria relevant to screening the full text of records.

EPA should clearly outline the process for handling anticipated overlap with literature relevant to multiple topics. EPA should describe how this will be addressed by the screeners and whether the same reviewer will be responsible for screening papers with inclusion/exclusion criteria across multiple topics or whether different reviewers are responsible only for screening studies for one particular topic. Additional details in regards to the process by which this screening will occur would be helpful. Given the breadth of each assessment (searching literature related to fate, engineering, exposure, human health, and environmental hazard) and the complexity of the screening process (tagging *on-topic* and *off-topic* literature and using additional sub-categories or sub-tags to allow for additional categorization), there appears to be the potential for individual papers to fall into different topic categories and have many different tags and sub-tags applied to indicate their relevance. However, it is unclear how this will be organized in the screening phase. Search strategies and inclusion/exclusion criteria appear to have been developed specifically for each literature topic and the potential overlap of literature relevant to multiple topics is not addressed.

EPA should explicitly include stopping rules for when the list of relevant studies will be considered final. There is no discussion of stopping dates or the process of updating the literature search to search for newer studies. Newer scientific studies will inevitably continue to appear in scientific journals and it will be impossible to continually attempt to include all these studies in a chemical assessment. To meet the deadlines as mandated by the Lautenberg Amendments, EPA should state clear stopping rules in the form of deadlines or criteria for when the body of included relevant studies will be finalized for the purposes of the chemicals assessment.

EPA should ensure gray literature search results are adequately screened. EPA's gray literature search strategy proposes to utilize Google's API to develop custom searches and return the first 100 results, sorted by predicted relevancy so that the results likely to be most relevant are screened first. It is unclear why this number is limited to only the first 100 and whether there was an empirical reason for why this particular number was selected. We recommend that EPA ensure that an adequate number of search results are screened, in particular considering that the gray literature can contribute potentially important information relevant to toxicity, mode of action, exposure, fate and transport, engineering or occupational exposure, or existence of publication bias.

EPA should consider "snowball searching," where the citations of included (i.e., *on-topic*) references are searched as well as using databases such as Web of Science to search for references that cite the included citations. EPA states that it plans on assessing the specificity and efficiency of the literature searches, through comparison of references either cited in existing problem formulation and risk assessment documents, in the public use documents and supporting life cycle diagrams, and comparison

of the references cited in review articles. Snowball searching will contribute to the evaluation of the specificity and efficiency of its literature searches, and also help to identify newer relevant studies that could potentially be included that have not yet been indexed in main databases such as PubMed.

EPA should incorporate appropriate tools for updating and evaluating systematic reviews in their chemical assessments. Garner et al., as part of efforts by a Cochrane Collaboration panel for updating guidance for systematic reviews, published guidance in 2016 for determining when it is appropriate to update a systematic review and outlining the steps for performing the update.[8] We have attached this guidance as an Appendix to these comments. EPA should evaluate the Cochrane tool's applicability to environmental chemicals given that Cochrane systematic reviews are geared towards reviews of clinical intervention evidence, so these tools may require updating and tailoring for an application to environmental health data.

A recent NAS report[9] recommends EPA should develop policies and procedures to allow the agency to identify, use and update existing systematic reviews. The committee also noted that it was important that the existing systematic review's study question directly addresses EPA's topic of interest and that the methods are critically evaluated before the systematic review is used and updated. EPA should ensure that only the highest quality systematic reviews be considered appropriate for use. It will be critical for EPA to develop tools to assist with the process of evaluating existing systematic reviews, particularly as this field continues to rapidly expand and more systematic reviews relevant to environmental health questions are published in the scientific literature, potentially of variable quality.

One tool which might be helpful for evaluating the risk of bias in systematic reviews is the ROBIS tool, which the NAS committee utilized in their report.[10] This tool was developed using rigorous methodology and can be applied for evaluating internal validity of systematic reviews in conjunction with other available tools to critically appraise and assess their quality. Of particular note is the strong emphasis on the recommendation that tools such as ROBIS should not be used to generate a composite quality score, as it has been well-documented that scoring can lead to bias in evaluation of the studies.[11] As such, the ROBIS tool presents several options for visually and graphically presenting results from risk of bias assessments based on individual domains or the overall rating, enabling reviewers to highlight particular areas of concern or reviews that are most relevant to the target question of interest.

Another tool which may be helpful in this process is the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), used by authors of systematic reviews to improve the reporting

---

[8] Garner P, Hopewell S, Chandler J, MacLehose H, Akl EA, Beyene J, Chang S, Churchill R, Dearness K, Guyatt G, Lefebvre C. When and how to update systematic reviews: consensus and checklist. bmj. 2016 Jul 20;354:i3507.

[9] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[10] Id.

[11] Jüni, P., Witschi, A., Bloch, R. and Egger, M., 1999. The hazards of scoring the quality of clinical trials for meta-analysis. *Jama*, *282*(11), pp.1054-1060.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

ED_004886_00001537-00008

of elements relevant to the systematic review and meta-analyses.[12] Increasingly, scientific journals are requiring the inclusion of checklists such as PRISMA with the submission of systematic reviews considered for publication. Although this tool is used to evaluate study reporting, and is not an assessment instrument to gauge the quality of a systematic review, it can still provide a useful framework to identify reported components of an existing systematic review in the process of evaluating quality or to identify missing components requiring follow-up with study authors to obtain additional information.

Furthermore, we strongly encourage EPA to evaluate the potential for financial conflicts of interest as an element in their study design. This is currently included as a consideration in evaluation risks of bias in some frameworks, such as the Navigation Guide,[13] and extracted for consideration as an additional domain in other frameworks, such as that developed by the National Toxicology Program's (NTP) Office of Health Assessment and Translation (OHAT).[14] The Cochrane Collaboration's risk of bias tool does not currently include a specific domain for bias related to study funding source, but this is an area of active discussion among its members.[15] The Cochrane Collaboration has recognized the importance of identifying study funding source, which has been empirically shown to be associated with biases.[16] A recent report from the NRC recommended that the U.S. EPA consider funding sources in their risk of bias assessment conducted for systematic reviews.[17]

We also strongly recommend EPA identify tools that may potentially not be appropriate for human health chemical assessments. Many tools are currently being developed for evaluating risk of bias, quality, and strength of evidence for individual studies as well as for systematic reviews. It is critical that EPA evaluate tools developed in other fields that may be relevant, such as for clinical or preclinical animal or human studies, as these tools could potentially be modified for an application to questions of environmental health relevance. However, these tools should be applied with caution—due to the

---

[12] http://www.prisma-statement.org/; Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic reviews. 2015 Jan 1;4(1):1.

[13] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

[14] NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[15] Bero LA. Why the Cochrane risk of bias tool should include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000075.*

Sterne JA. Why the Cochrane risk of bias tool should not include funding source as a standard item. *Cochrane Database Syst Rev. 2013 Dec 20; (12):ED000076.*

[16] Krauth D, Anglemyer A, Philipps R, Bero L. Nonindustry-sponsored preclinical studies on statins yield greater efficacy estimates than industry-sponsored studies: a meta-analysis. *PLoS Biol. 2014 Jan; 12(1):e1001770*

Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev. 2012 Dec 12; 12():MR000033.*

[17] NRC (National Research Council). Review of EPA's Integrated Risk Information System (IRIS) Process. Washington, DC:National Academies Press. 2014.

ED_004886_00001537-00009

differences in the types of evidence under evaluation a direct application to a difference evidence base than intended could lead to biased and erroneous conclusions.[18]

2. **EPA needs to consider aggregate exposure within and across populations; otherwise it will underestimate risk. Aggregate exposure should include legacy uses, uses where a chemical is present as a contaminant or by-product, and uses already assessed by EPA.**

In general, EPA is proposing to consider three populations for exposure assessment: 1) Occupational users and non-users; 2) consumers and bystanders; and 3) general population. We strongly recommend that EPA calculate the aggregate exposures within and across these populations-- risk will be underestimated if it does not include these real-world exposures. Exposures within a population should also be aggregated (rather than considered in isolation) in order to estimate the general population's actual exposure to the chemical—for example, through exposures from food, water and air.

Further, as shown in the Figure below, exposures must also be aggregated across populations. Consumers and workers are part of the general population – that is, since workers and consumers also eat food and drink water, they will have the same exposures as the general population, in addition to the anticipated exposures on-the-job or from consumer products. Some workers will also be consumer product users, so they have the potential to face general, consumer product, and on-the-job exposures. These specific exposure scenarios must be accounted for in EPA's exposure estimation to ensure that such individual exposures are adequately considered and integrated into the risk assessment.

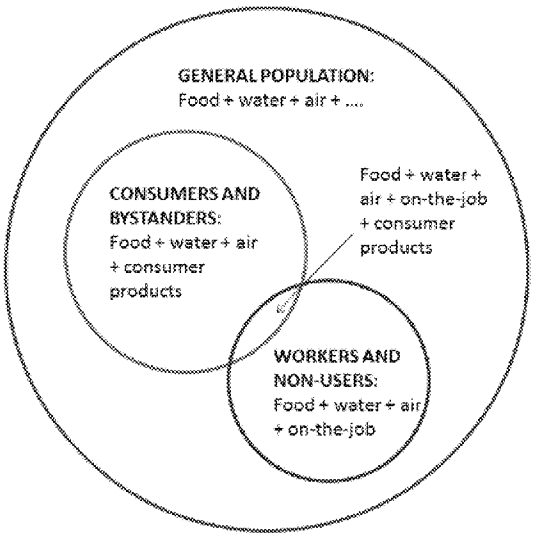


Figure: EPA must assess aggregate exposures within and across all the populations for accurate exposure assessment.

In the Introduction section of the chemical Scope documents, EPA states that it "may consider background exposures from legacy use, associated disposal, and legacy disposal as part of an assessment of aggregate exposure or as a tool to evaluate the risk of exposures resulting from non-

[18] Nachman, K.E., Lam, J., Schinasi, L.H., Smith, T.C., Feingold, B.J. and Casey, J.A., 2017. O'Connor et al. systematic review regarding animal feeding operations and public health: critical flaws may compromise conclusions. *Systematic Reviews*, *6*(1), p.179.

legacy uses."[19] This falls short of the analysis required under Lautenberg TSCA. It is critical that EPA consider ongoing exposures from legacy uses and disposal, and includes these as part of the aggregate exposure assessment. Asbestos and HBCD are two examples of this, as they have enormous volumes in place in buildings and existing infrastructure. The Healthy Building Network estimates there are 66 million- 132 million pounds (30,000-60,000 metric tons) of HBCD in insulation in existing buildings[20] — these reservoirs in-place are and will continue to be critical sources of ongoing exposures. HBCD was also used in cars and furniture, which are long-lived consumer items that will continue to contribute to ongoing exposures for years to come.

Another example is 1,4-dioxane, which was historically used as a chemical stabilizer for chlorinated solvents. Many groundwater aquifers are contaminated with 1,4-dioxane, and the extent of legacy contamination of groundwater is likely underestimated.[21] Also, 1,4-dioxane occurs in a wide variety of products including personal care products, detergents, waxes, and antifreeze, and 1,4-dioxane is a by-product in manufacturing processes involving ethylene oxide, such as the production of polyethylene terephthalate (PET), polyester, and surfactants. The use and disposal of 1,4-dioxane has led to past environmental contamination which contributes to on-going exposures.[22] The physical and chemical properties of 1,4-dioxane render it a persistent and highly mobile water contaminant: it is highly miscible in water.[23] Exposures via drinking water are documented back to the 1980s and continue today.[24] Results from EPA's Third Unregulated Contaminant Monitoring Rule (UCMR3) highlight that over 13% of 4,905 public drinking water systems serving >10,000 people had concentrations of 1,4-dioxane above the EPA Reference Concentration of 0.35 ppb 1,4-dioxane.[25] Furthermore, the UCMR3 results do not capture exposures in communities served by small public drinking water systems serving <10,000 people.[26] Approximately 27% of the US population is served by small public drinking water systems.[27] Thus, it will be critical for EPA to consider the population's current exposure to 1,4-dioxane via sources like drinking water as part of their assessment for health risks.

---

[19] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 12

[20] Safer Chemicals, Healthy Families et al. Comments to the U.S. Environmental Protection Agency (EPA) on the Scope of its Risk Evaluation for the TSCA Work Plan Chemicals: CYCLIC ALIPHATIC BROMIDE CLUSTER or HEXABROMOCYCLODODECANE (HBCD). March 15, 2017. https://healthybuilding.net/uploads/files/saferchemicals-hbcd.pdf

[21] David T. Adamson, Shaily Mahendra, Kenneth L. Walker, Jr., Sharon R. Rauch, Shayak Sengupta, and Charles J. Newell. A Multisite Survey To Identify the Scale of the 1,4-Dioxane Problem at Contaminated Groundwater Sites. *Environmental Science & Technology Letters* 2014 1 (5), 254-258 DOI: 10.1021/ez500092u

[22] Agency for Toxic Substances and Disease Registry. TOXICOLOGICAL PROFILE FOR 1,4-DIOXANE. Atlanta, GA; 2012.

[23] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

[24] Dietrich, A.M., D.S. Millington, and R.F. Christman. 1983. Specific identification of organic pollutants in Haw River water using gas chromatography/mass spectrometry. WRRI Report No. 206. North Carolina State University, Raleigh, NC.

[25] U.S. Environmental Protection Agency. Integrated Risk Information System (IRIS). 1,4-Dioxane (CASRN 123-91-1). Washington, D.C.; 2013. http://www.epa.gov/iris/subst/0326.htm.

U.S. Environmental Protection Agency. Unregulated Contaminant Monitoring Rule 74 FR 51850, EPA-HQ-OW-2007-1189 FRL-8963-6. Washington, D.C.; 2009. https://federalregister.gov/a/E9-24287.s

[26] Knappe, D.R.U., Lopez-Velandia, C., Hopkins, Z., & Sun, M. (2016). Occurrence of 1,4-Dioxane in the Cape Fear River Watershed and Effectiveness of Water Treatment Options for 1,4-Dioxane Control. NC Water Resources Research Institute.

[27] EPA. 2017. Small Drinking Water Systems Research https://www.epa.gov/water-research/small-drinking-water-systems-research-0 (accessed September 17, 2017).

ED_004886_00001537-00011

When a chemical is present in products or media as a contaminant/ by-product, EPA needs to include and assess these exposures. We strongly recommend against ignoring or discounting these potential exposures routes. For example, EPA proposes to exclude from consideration conditions of use of 1,4-dioxane when it is present as contaminant in a wide variety of items, including household detergents, cosmetics/ toiletries, and foods.[28] This exclusion is not scientifically justified. Cosmetics and personal care products have the potential to contribute significantly to exposures, since people are applying them directly to their bodies, often multiple times per day, every day.

Finally, in the exposure assessments for methylene chloride, N-methylpyrrolidone and trichloroethylene, EPA is proposing to exclude uses it already assessed.[29] We agree that EPA does not need to re-assess these uses; these evaluations have been completed and finalized. However, unless and until such uses are banned, the exposures from these uses continue. Therefore, the new risk evaluations need to consider the contributions of these uses to exposures by using the exposure values from the previous assessments.

For the occupational exposure analysis plan, EPA states it will "Consider and incorporate applicable engineering controls and/or personal protective equipment into exposure scenarios."[30] However, these are not realistic assumptions nor are they appropriate for public health protection. EPA's own research shows that the primary factors influencing whether a user understands label information are the users' literacy and numeracy, which frequently correlate with the users' education and income.[31] Therefore, people with less education, lower income, and less advanced literary skills will be the most likely to not understand label instructions. These individuals already disproportionately bear the burden of exposures to multiple environmental hazards and the resulting health impacts; thereby placing further burden on this already stressed susceptible subpopulation. Further, appropriate personal protective equipment (PPE) for workers is often not provided by employers, or may not be fitted or working properly. When evaluating occupational exposures, EPA needs to take into consideration all potential and feasible routes of exposure, and should not exclude exposure routes based on assumptions of PPE and/ or exposure controls in place. These controls are not guaranteed and may change in the future, so to assume zero exposure via these routes would be inappropriate and a failure to adequately ensure health protections, especially for susceptible sub-populations as required by the Lautenberg TSCA.

In summary, EPA needs to account for all the sources of exposure or it will underestimate risk for all 10 chemicals. When analyzing aggregate exposures, "sentinel exposure" may be considered simultaneously, where appropriate. However, these are not mutually exclusive and EPA should not incorporate sentinel to the exclusion of aggregate.

3. EPA appropriately identifies factors to consider to identify populations subject to greater exposures. EPA should also address susceptible sub-populations, following recommendations

---

[28] US EPA (2017). Scope of the Risk Evaluation for 1,4-dioxane. Pg. 21

[29] US EPA (2017). Scope of the Risk Evaluation for Methylene Chloride (Dichloromethane, DCM). Pg. 30
US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 19-20
US EPA (2017). Scope of the Risk Evaluation for Trichloroethylene. Pg. 27

[30] See, for example, US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg. 45

[31] US Environmental Protection Agency (EPA). 2016. "The Effectiveness of Labeling on Hazardous Chemicals and Other Products." Office of Chemical Safety and Pollution Prevention. RIN 2079-AK07.

ED_004886_00001537-00012

from the National Academies of Sciences (NAS) to identify susceptible sub-populations based on established extrinsic and intrinsic factors that increase vulnerability.

In general, EPA proposes to consider workers and occupational non-users, consumer and by-standers, and other groups within the general population in proximity to conditions of use as sub-populations who experience greater exposures. In particular, EPA has appropriately identified people who live or work near manufacturing, processing, distribution, use or disposal sites as facing greater exposures. Such communities are often low income and/ or people of color, exposed to a disproportionate share of pollution, environmental hazards, social and economic stressors. Multiple exposures to chemical and non-chemical stressors collectively increase the risk of harm, combined with synergistic effects with other health stressors in their daily lives such as limited access to quality health care.[32,33]

EPA's risk evaluation needs to fully account for the reality of cumulative exposures, as recommended by the NAS in their Phthalates and Cumulative Risk report.[34] As described below, EPA can use "default values" to account for cumulative exposures.

In regards to greater susceptibility, EPA's considerations for addressing susceptibility vary considerably across the 10 chemicals. EPA should apply a consistent approach to addressing susceptibility across the 10 chemicals. The following are well-known factors that increase biologic sensitivity or reduce resilience to exposures,[35,36] and these as well as other relevant factors should be standard considerations for all 10 chemicals to identify susceptible sub-populations:

Intrinsic/ endogenous factors

- Genetic polymorphisms/ genetics/ genetic makeup
- Health status/ nutritional status/ disease status/ pre-existing conditions
- Prenatal lifestage
- Age

Extrinsic factors

- Multiple exposures/ co-exposures
- Race/ ethnicity
- Socioeconomic status (SES)

For example, the prenatal lifestage is the most sensitive to developmental and reproductive toxicants, and women of child-bearing age should be considered as a susceptible sub-population for any chemicals with such hazards. Women of reproductive age are not specifically identified as a potential susceptible

---

[32] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[33] Vesterinen HM, Morello-Frosch R, Sen S, Zeise L, Woodruff TJ. Cumulative effects of prenatal-exposure to exogenous chemicals and psychosocial stress on fetal growth: Systematic-review of the human and animal evidence. Meliker J, editor. PLoS One. 2017 Jul 12;12(7):e0176331.

[34] National Research Council. Committee on the Health Risks of Phthalates, Board on Environmental Studies and Toxicology, Division on Earth and Life Studies. 2008. Phthalates and cumulative risk assessment: the task ahead. Washington, D.C.: National Academies Press.

[35] Morello-Frosch R, Zuk M, Jerrett M, Shamasunder B, Kyle AD. Understanding the cumulative impacts of inequalities in environmental health: Implications for policy. Health Aff. 2011;30(5):879–87.

[36] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

sub-population for pigment violet 29, TCE, NMP, PERC, or HBCD, even though EPA will consider reproductive and developmental toxicity hazards for these chemicals.

As discussed below, science-based defaults should be used to account for these and other susceptibilities, unless there is there is chemical-specific data available to support increasing or decreasing the default.

4. **EPA should rely on existing IRIS assessments for hazard identification. Moving forward, EPA should complete hazard identification or add additional studies only through a systematic review process, which integrates animal, human and mechanistic evidence as recommended by the recent NAS report.**

EPA cites existing IRIS assessments for five chemicals; because these are EPA's own assessments which have gone through the Agency's peer-review process, and in some cases NAS review, EPA can rely on these existing finalized, authoritative assessments for hazard identification.

Moving forward, a weight of evidence evaluation is required by law, which EPA defines as:

> "Weight of scientific evidence means a systematic review method, applied in a manner suited to the nature of the evidence or decision, that uses a pre-established protocol to comprehensively, objectively, transparently, and consistently, identify and evaluate each stream of evidence, including strengths, limitations, and relevance of each study and to integrate evidence as necessary and appropriate based upon strengths, limitations, and relevance."[37]

Therefore, EPA should use a systematic review process for evaluating scientific information for chemicals that do not have an IRIS assessment and for any additional studies that will be considered for the chemicals that have IRIS assessments.

For the scoping document, EPA should include all hazards identified in the literature, and not make decisions about their relevance to the risk evaluation until a systematic review has been completed. For a number of chemicals, EPA has inappropriately drawn conclusions about hazards prior to the completion of a systematic review. Some examples are given in the table below where EPA concludes that HBCD, NMP and pigment violet 29 are not genotoxic based on previous assessments and without conducting a systematic review.

Chemical	Example Text from EPA Scoping Document
HBCD	"Available data suggest that HBCD is not genotoxic. Existing assessments have also concluded, based on genotoxicity information and a limited lifetime study, that HBCD is not carcinogenic (NICNAS, 2012; EINECS, 2008; TemaNord, 2008; OECD, 2007). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity or cancer hazards in the risk evaluation of HBCD at this time."[38]
NMP	"NMP is not mutagenic, based on results from bacterial and mammalian *in vitro* tests and *in vivo* systems and is not considered to be carcinogenic (RIVM, 2013; OECD,

[37] 82 Fed. Reg. 138, 33748
[38] US EPA (2017). Scope of the Risk Evaluation for Cyclic Aliphatic Bromides Cluster. Pg.36

	2007; WHO, 2001). Unless new information indicates otherwise, EPA does not expect to conduct additional in-depth analysis of genotoxicity and cancer hazards in the NMP risk evaluation."[39]
Pigment violet 29	"Testing for carcinogenicity of Pigment Violet 29 has not been conducted. However, negative genotoxicity results, structure-activity considerations and the expectation of negligible absorption and uptake of Pigment Violet 29 (based on very low solubility), indicate carcinogenicity of Pigment Violet 29 is unlikely. Unless new information indicates otherwise, EPA does not expect to conduct additional, in-depth analyses of genotoxicity and cancer hazards in the risk evaluation of Pigment Violet 29."[40]

The National Academies recently released a report with recommendations on implementation of systematic review for EPA's chemical evaluations (which we will refer to as the 'NAS Systematic Review report' for simplicity).[41] First, they recommend that EPA should develop policies and procedures that allow the agency to use and update existing systematic reviews, since the committee concluded that could potentially save time and resources. EPA should conduct a review to determine whether there are existing systematic reviews on the topic of interest and if there is, EPA should evaluate it to determine if it is high-quality. The NAS recommends that EPA should build on existing high-quality reviews to incorporate new studies and use the updated systematic review as a basis for its assessment. The assessments cited by EPA to support the hazard identification claims are not systematic reviews; even if they were, EPA should evaluate them for quality before relying on their conclusions.

Second, it is very likely that additional studies have been published since the assessments EPA cites were completed. EPA should develop criteria to evaluate the internal validity (risk of bias) of individual studies, utilizing existing tools that have been developed and empirically demonstrated on environmental health studies such as the Navigation Guide or the OHAT approach.[42] We also recommend that EPA not using a scoring system to evaluate study quality. Specifically, we note that empirically validated approaches in the clinical sciences such as Cochrane discourage using a numerical scale scoring approach for evaluating study quality because calculating a score requires choosing a weighting scheme for each component, which generally is nearly impossible to justify.[43] Furthermore, a study might be well designed to eliminate bias, but because the study failed to report details in the publication under review, it will receive a low score--most available scoring systems include a mix of risk of bias and reporting biases which is inappropriate. Additionally, quality scores have been shown to be invalid for assessing risk of bias in clinical research.[44] The current standard in evaluation of both clinical

[39] US EPA (2017). Scope of the Risk Evaluation for N-Methylpyrrolidone. Pg. 36

[40] US EPA (2017). Scope of the Risk Evaluation for Pigment Violet 29. Pg.29

[41] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017.

[42] Woodruff, T.J. and Sutton, P., 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health affairs*, *30*(5), pp.931-937.

Woodruff, T.J. and Sutton, P., 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. *Environmental health perspectives*, *122*(10), p.1007.

NTP OHAT Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookjan2015_508.pdf

[43] Juni, P., A. Witschi, R. Bloch, and M. Egger. 1999. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 282(11):1054-1060.

[44] Id.

ED_004886_00001537-00015

and environmental health research calls for reporting each component of the assessment tool separately and not calculating an overall numeric score.[45]

Data generated by alternative test methods (such as high-throughput screening methods) are not different than any other type of *in vitro* or cell-based assay data that would be considered in a systematic review. These kinds of assays provide mechanistic data, and the NAS Systematic Review report explicitly considered how mechanistic data could be utilized in a systematic review for evidence integration. The committee came to two conclusions. First, the same protocol for evaluating relevance and study quality must be used with mechanistic data as for any other study. For example, in the report's case study on phthalates, the committee was not able to integrate results from high-throughput assays because the cell lines used were of unknown relevance to the *in vivo* mechanism of phthalate toxicity.[46] Second, the foundation of the hazard classification in a systematic review is the animal and human data, with the mechanistic data playing a supporting role. If mechanistic data is relevant, it can be used to upgrade a hazard classification, or increase the confidence of a finding made based on evaluation of animal and human data. A hazard classification is never made based on high-throughput or other kinds of mechanistic data alone.[47]

5. For risk characterization, EPA should use defaults and methods that account for the full range of risks in the population and that will form the basis of decisions that protect the public's health.

Defaults

We strongly support the use of health protective defaults to incorporate factors that reflect the range of variability and susceptibility in the population to ensure risks are not underestimated. The importance of using protective science-based defaults was highlighted by the NAS in 2009.[48] The default should be used for factors that are known to influence risk unless there is chemical-specific data that support increasing or decreasing it; when there is inadequate information to quantitatively assess inter- or intra-species differences for a specific chemical, the defaults should be used. For example, EPA's defaults should include:

- Inter-human variability, general
- Inter-human susceptibility to carcinogens, adult
- Inter-human susceptibility to carcinogens, early life (including prenatal)
- Inter-human susceptibility to non-carcinogens, early life (including prenatal)
- Animal findings are relevant to humans
- Findings from one route of exposure are considered representative unless data show otherwise

---

[45] Higgins, J.P.T., and S. Green, eds. 2008. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons.

Whiting, P., Harbord, R. and Kleijnen, J., 2005. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, *5*(1), p.19.

Whiting, P., Savović, J., Higgins, J.P., Caldwell, D.M., Reeves, B.C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R., 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, *69*, pp.225-234.

[46] The National Academies of Sciences. Application of Systematic Review Methods in an Overall Strategy for Evaluating Low-Dose Toxicity from Endocrine Active Chemicals. Washington, D.C.: National Academies Press; 2017. Pg. 78

[47] Id. Pp. 158-9

[48] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch 4-6

ED_004886_00001537-00016

EPA has relied on standard default values ("uncertainty" or "safety" factors) that have been applied across the board to various chemicals and health outcomes. But newer science demonstrates that EPA's typical safety factor of 10 is insufficient to account for variability due to life stage, genetics, underlying disease status, and external stressors that may be due to poverty or other difficult life conditions.

For cancer, the NAS recommended that EPA include a factor to account for human variability in response to carcinogens, as EPA's current approach inaccurately assumes that there is no variability in response. They found that a factor of 25- to 50- may account for the variability between the median individual and those with more extreme responses, and recommended 25 as a reasonable default value.[49]

Similarly, EPA should increase or add factors that address cancer and non-cancer susceptibility during early life stages. While EPA does account for increased susceptibility to genotoxicants, it does not include the prenatal period or chemicals that can influence cancer through other mechanisms. California EPA's guidance incorporates factors to account for increased susceptibility for exposures that occur prenatally for carcinogens, non-mutagenic carcinogenic agents and non-carcinogens. Their literature review on differential susceptibility to carcinogens and non-carcinogens based on age and life stage derived age adjustment values for carcinogens which include the prenatal period[50] and increased the default intraspecies uncertainty factors for non-carcinogens to 30 and 100 for specific endpoints such as asthma or neurotoxicity.[51] At a minimum, EPA should use Cal EPA's age adjustment values and intraspecies uncertainty factors for incorporating age/early life susceptibility.

In general, developmental life stages, including the fetus, infancy, and childhood, are more vulnerable to chemical exposure and toxicity. However, typical EPA age-dependent adjustment factors account for other life stages but NOT fetal exposures. Recent studies have demonstrated differential expression and activity of metabolic enzymes such as Cytochrome P450 in fetal versus adult tissue, indicating potential lifestage-dependent variability in metabolic capabilities and greater vulnerability during fetal development not accounted for in current risk assessment practices.[52] This is a critical point to address, as disruptions during fetal development have implications for health and disease in adulthood. EPA should evaluate this rich body of literature to identify the most up-to-date scientific knowledge regarding human variability and susceptibility and incorporate these scientifically-based default values in their assessments unless there are chemical-specific data supporting departing from the defaults. California EPA also developed child-specific risk values for chemicals (e.g., atrazine, lead, nickel, manganese, heptachlor) that specifically address routes of exposure and differences in susceptibility

---

[49] Id. Pg. 168

[50] California EPA 2009. Cal EPA 2009. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document for Cancer Potency Factors: Methodologies for derivation, listing of available values, and adjustments to allow for early life stage exposures. http://oehha.ca.gov/media/downloads/crnr/tsdcancerpotency.pdf

[51] Cal EPA 2008. California Environmental Protection Agency, Office of Environmental Health Hazard Assessment. Technical Support Document For the Derivation of Noncancer Reference Exposure Levels http://oehha.ca.gov/media/downloads/crnr/noncancertsdfinal.pdf

[52] Sadler, N.C., Nandhikonda, P., Webb-Robertson, B.J., Ansong, C., Anderson, L.N., Smith, J.N., Corley, R.A. and Wright, A.T., 2016. Hepatic cytochrome P450 activity, abundance, and expression throughout human development. *Drug Metabolism and Disposition*, *44*(7), pp.984-991.

ED_004886_00001537-00017

unique to children compared to adults.[53] EPA should review these evaluations and incorporate these values as appropriate. Furthermore, a default guidance principle should be that animal findings are relevant to humans unless there is sufficient and compelling information to support otherwise.

Risk estimates

EPA should not use MOE (margin of exposure) as an analysis method in the risk evaluation process moving forward. MOE is not an estimate of risk—it is a single number that is a version of the "bright line" approach like the Reference Dose (or Reference Concentration for inhalation doses). MOE is calculated by dividing the point of departure (e.g., LOAELs, NOAELs or BMDLs) by estimated exposure values, and this 'bright line' approach does not provide information about the magnitude of the risks above, at, or below this line. Further, it implies that there is a "safe" level of exposure below which no harm will occur. While this may be true for a select few chemicals, the NAS Science and Decisions report recognizes that this is not a valid assumption for all chemicals and has recommended moving away from such "bright line" approaches which do not establish risk estimates across the full range of exposures.[54] Additionally, the MOE will not provide the necessary information for future analysis of risks and benefits that will be critical for decision-making on these chemicals.[55] We recommend that EPA utilize available analytical methods such as PODs based on a BMD to develop quantified estimates of risk.

EPA appropriately states that a dose-response assessment will be conducted for all identified human health hazard endpoints. PODs should also be developed for every endpoint unless the data are insufficient to develop a model. For calculating cancer or non-cancer risks, we recommend always using a point of departure (POD) of a benchmark dose (BMD) at 1%. The POD should be based on a BMD calculation, not the NOAEL/LOAEL, unless the data are insufficient to model. EPA already recognizes the features that make BMDs superior: BMDs account for the shape of the dose–response function; are independent of study design, such as the space between dosing; and are comparable across chemicals.[56]

Historically, for carcinogens that are direct mutagens or are associated with large human body burdens, EPA has assumed there is no threshold of effect. But the NAS Science and Decisions report highlights the science indicating that this linear presumption with no threshold is appropriate for the calculation of both cancer and non-cancer risks, and regardless of whether a carcinogen is a mutagen. For example, dose-response relationships can be linear at low dose when exposures contribute to an existing disease process, add to background processes and/ or exposures, and interact with interindividual variability or susceptibility.[57] Science and Decisions recommends harmonizing cancer and non-cancer risk assessment approaches. Therefore, for calculating non-mutagen cancer or non-cancer risks based on a POD, EPA should use the same approach as for mutagens, which assumes a straight line from the POD. In fact, a

---

[53] California Environmental Protection Agency. Office of Environmental Health Hazard Assessment (OEHHA). Child-Specific Reference Doses (chRDs) Finalized to Date. Available from: http://oehha.ca.gov/risk-assessment/chrd/table-all-chrds

[54] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009.

[55] McGartland A, Revesz R, Axelrad DA, Dockins C, Sutton P, Woodruff TJ. Estimating the health benefits of environmental regulations. Science (80-). 2017 Aug 4;357(6350):457–8.

[56] Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, Rusyn I. 2014. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. Environmental health perspectives. 122(5).

[57] National Research Council. Science and Decisions: Advancing Risk Assessment. Washington, D.C.: National Academies Press; 2009. Ch.5

ED_004886_00001537-00018

linear relationship may actually underestimate risks for some chemicals where the dose-response curve is supra-linear.

6. Confidential Business Information (CBI) claims should not be used to obscure critical data and information from the public.

Production volumes for both asbestos and pigment violet 29 have been claimed as CBI. Production volume is basic information about a chemical to which the public and scientists should have access. We urge EPA to move forward with substantiating such claims under the new TSCA.

Appendix

When and how to update systematic reviews: consensus and checklist
Garner et al.
BMJ 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

ED_004886_00001537-00020

OPEN ACCESS

When and how to update systematic reviews: consensus and checklist

Paul Garner,[1] Sally Hopewell,[2] Jackie Chandler,[3] Harriet MacLehose,[3] Elie A Akl,[5,6] Joseph Beyene,[7] Stephanie Chang,[8] Rachel Churchill,[9] Karin Dearness,[10] Gordon Guyatt,[4] Carol Lefebvre,[11] Beth Liles,[12] Rachel Marshall,[3] Laura Martínez García,[13] Chris Mavergames,[14] Mona Nasser,[15] Amir Qaseem,[16,17] Margaret Sampson,[18] Karla Soares-Weiser,[3] Yemisi Takwoingi,[19] Lehana Thabane,[4,20] Marialena Trivella,[21] Peter Tugwell,[22] Emma Welsh,[23] Ed C Wilson,[2,4] Holger J Schünemann[4,5]

For numbered affiliations see end of article.

Correspondence to: P Garner Paul.Garner@lstmed.ac.uk

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2016;354:i3507
http://dx.doi.org/10.1136/bmj.i3507

Accepted: 26 May 2016

Updating of systematic reviews is generally more efficient than starting all over again when new evidence emerges, but to date there has been no clear guidance on how to do this. This guidance helps authors of systematic reviews, commissioners, and editors decide when to update a systematic review, and then how to go about updating the review.

Systematic reviews synthesise relevant research around a particular question. Preparing a systematic review is time and resource consuming, and provides a snapshot of knowledge at the time of incorporation of data from studies identified during the latest search. Newly identified studies can change the conclusion of a review. If they have not been included, this threatens the validity of the review, and, at worst, means the review could mislead. For patients and other healthcare consumers, this means that care and policy development might not be fully informed by the latest research; furthermore, researchers could be misled and carry out research in areas where no further research is actually needed.[1] Thus, there are clear benefits to updating reviews, rather than duplicating the entire process as new evidence emerges or new methods develop. Indeed, there is probably added value to updating a review, because this will include taking into account comments and criticisms, and adoption of new methods in an iterative process.[2-6]

Cochrane has over 20 years of experience with preparing and updating systematic reviews, with the publication of over 6000 systematic reviews. However, Cochrane's principle of keeping all reviews up to date has not been possible, and the organisation has had to adapt: from updating when new evidence becomes available,[7] to updating every two years,[8] to updating based on need and priority.[9] This experience has shown that it is not possible, sensible, or feasible

SUMMARY POINTS

Updating systematic reviews is, in general, more efficient than starting afresh when new evidence emerges. The panel for updating guidance for systematic reviews (PUGs; comprising review authors, editors, statisticians, information specialists, related methodologists, and guideline developers) met to develop guidance for people considering updating systematic reviews. The panel proposed the following:

1. Decisions about whether and when to update a systematic review are judgments made for individual reviews at a particular time. These decisions can be made by agencies responsible for systematic review portfolios, journal editors with systematic review update services, or author teams considering embarking on an update of a review.
2. The decision needs to take into account whether the review addresses a current question, uses valid methods, and is well conducted; and whether there are new relevant methods, new studies, or new information on existing included studies. Given this information, the agency, editors, or authors need to judge whether the update will influence the review findings or credibility sufficiently to justify the effort in updating it.
3. Review authors and commissioners can use a decision framework and checklist to navigate and report these decisions with "update status" and rationale for this status. The panel noted that the incorporation of new synthesis methods (such as Grading of Recommendations Assessment, Development and Evaluation (GRADE)) is also often likely to improve the quality of the analysis and the clarity of the findings.
4. Given a decision to update, the process needs to start with an appraisal and revision of the background, question, inclusion criteria, and methods of the existing review.
5. Search strategies should be refined, taking into account changes in the question or inclusion criteria. An analysis of yield from the previous edition, in relation to databases searched, terms, and languages can make searches more specific and efficient.
6. In many instances, an update represents a new edition of the review, and authorship of the new version needs to follow criteria of the International Committee of Medical Journal Editors (ICMJE). New approaches to publishing licences could help new authors build on and re-use the previous edition while giving appropriate credit to the previous authors.

The panel also reflected on this guidance in the context of emerging technological advances in software, information retrieval, and electronic linkage and mining. With good synthesis and technology partnerships, these advances could revolutionise the efficiency of updating in the coming years.

ED_004886_00001537-00021

to continually update all reviews all the time. Other groups, including guideline developers and journal editors, adopt updating principles (as applied, for example, by the *Systematic Reviews* journal; https://systematicreviewsjournal.biomedcentral.com/).

The panel for updating guidance for systematic reviews (PUGs) group met to draw together experiences and identify a common approach. The PUGs guidance can help individuals or academic teams working outside of a commissioning agency or Cochrane, who are considering writing a systematic review for a journal or to prepare for a research project. The guidance could also help these groups decide whether their effort is worthwhile.

Panel selection and procedures

An international panel of authors, editors, clinicians, statisticians, information specialists, other methodologists, and guideline developers was invited to a two day workshop at McMaster University, Hamilton, Canada, on 26-27 June 2014, organised by Cochrane. The organising committee selected the panel (web appendix 1). The organising committee invited participants, put forward the agenda, collected background materials and literature, and drafted the structure of the report.

The purpose of the workshop was to develop a common approach to updating systematic reviews, drawing on existing strategies, research, and experience of people working in this area. The selection of participants aimed on broad representation of different groups involved in producing systematic reviews (including authors, editors, statisticians, information specialists, and other methodologists), and those using the reviews (guideline developers and clinicians). Participants within these groups were selected on their expertise and experience in updating, in previous work developing methods to assess reviews, and because some were recognised for developing approaches within organisations to manage updating strategically. We sought to identify general approaches in this area, and not be specific to Cochrane; although inevitably most of the panel were somehow engaged in Cochrane.

The workshop structure followed a series of short presentations addressing key questions on whether, when, and how to update systematic reviews. The proceedings included the management of authorship and editorial decisions, and innovative and technological approaches. A series of small group discussions followed each question, deliberating content, and forming recommendations, as well as recognising uncertainties. Large group, round table discussions deliberated further these small group developments. Recommendations were presented to an invited forum of individuals with varying levels of expertise in systematic reviews from McMaster University (of over 40 people), widely known for its contributions to the field of research evidence synthesis. Their comments helped inform the emerging guidance.

The organising committee became the writing committee after the meeting. They developed the guidance arising from the meeting, developed the checklist and diagrams, added examples, and finalised the manuscript. The guidance was circulated to the larger group three times, with the PUGs panel providing extensive feedback. This feedback was all considered and carefully addressed by the writing committee. The writing committee provided the panel with the option of expressing any additional comments from the general or specific guidance in the report, and the option for registering their own view that might differ to the guidance formed and their view would be recorded in an annex. In the event, consensus was reached, and the annex was not required.

Definition of update

The PUGs panel defined an update of a systematic review as a new edition of a published systematic review with changes that can include new data, new methods, or new analyses to the previous edition. This expands on a previous definition of a systematic review update.[10] An update asks a similar question with regard to the participants, intervention, comparisons, and outcomes (PICO) and has similar objectives; thus it has similar inclusion criteria. These inclusion criteria can be modified in the light of developments within the topic area with new interventions, new standards, and new approaches. Updates will include a new search for potentially relevant studies and incorporate any eligible studies or data; and adjust the findings and conclusions as appropriate. Box 1 provides some examples.

Which systematic reviews should be updated and when?

Any group maintaining a portfolio of systematic reviews as part of their normative work, such as guidelines panels or Cochrane review groups, will need to prioritise which reviews to update. Box 2 presents the approaches used by the Agency for HealthCare Research and Quality (AHRQ) and Cochrane to prioritise which systematic reviews to update and when. Clearly, the responsibility for deciding which systematic reviews should be updated and when they will be updated will vary: it may be centrally organised and resourced, as with the AHRQ scientific resource centre (box 2). In Cochrane, the decision making process is decentralised to the Cochrane Review Group editorial team, with different approaches applied, often informally.

Box 1: Examples of what factors might change in an updated systematic review

- A systematic review of steroid treatment in tuberculosis meningitis used GRADE methods and split the composite outcome in the original review of death plus disability into its two components. This improved the clarity of the reviews findings in relation to the effects and the importance of the effects of steroids on death and on disability.[11]
- A systematic review of dihydroartemisinin-piperaquine (DHAP) for treating malaria was updated with much more detailed analysis of the adverse effect data from the existing trials as a result of questions raised by the European Medicines Agency. Because the original review included other comparisons, the update required extracting only the DHAP comparisons from the original review, and a modification of the title and the PICO.[12]
- A systematic review of atorvastatin was updated with simple uncontrolled studies.[13] This update allowed comparisons with trials and strengthened the review findings.[14]

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001537-00022

The PUGs panel recommended an individualised approach to updating, which used the procedures summarised in figure 1. The figure provides a status category, and some options for classifying reviews into each of these categories, and builds on a previous decision tool and earlier work developing an updating classification system.[15 16] We provide a narrative for each step.

Step 1: assess currency

Does the published review still address a current question?

An update is only worthwhile if the question is topical for decision making for practice, policy, or research priorities (fig 1). For agencies, people responsible for managing a portfolio of systematic reviews, there is a need to use both formal and informal horizon scanning. This type of scanning helps identify questions with currency, and can help identify those reviews that should be updated. The process could include monitoring policy debates around the review, media outlets, scientific (and professional) publications, and linking with guideline developers.

Has the review had good access or use?

Metrics for citations, article access and downloads, and sharing via social or traditional media can be used as proxy or indicators for currency and relevance of the review. Reviews that are widely cited and used could be important to update should the need arise. Comparable reviews that are never cited or rarely downloaded, for example, could indicate that they are not addressing a question that is valued, and might not be worth updating.

In most cases, updated reviews are most useful to stakeholders when there is new information or methods that result in a change in findings. However, there are some circumstances in which an up to date search for information is important for retaining the credibility of the review, regardless of whether the main findings would change or not. For example, key stakeholders would dismiss a review if a study is carried out in a relevant geographical setting but is not included; if a large, high profile study that might not change the findings is not included; or if an up to date search is required for a guideline to achieve credibility. Box 3 provides such examples. If the review does not answer a current question, the intervention has been superseded, then a decision can be made not to update and no further intelligence gathering is required (fig 1).

Did the review use valid methods and was it well conducted?

If the question is current and clearly defined, the systematic review needs to have used valid methods and be well conducted. If the review has vague inclusion criteria, poorly articulated outcomes, or inappropriate methods, then updating should not proceed. If the question is current, and the review has been cited or used, then it might be appropriate to simply start with a new protocol. The appraisal should take into account the methods in use when the review was done.

Step 2: identify relevant new methods, studies, and other information

Are there any new relevant methods?

If the question is current, but the review was done some years ago, the quality of the review might not meet current day standards. Methods have advanced quickly, and data extraction and understanding of the review process have become more sophisticated. For example:

- Methods for assessing risk of bias of randomised trials,[23] diagnostic test accuracy (QUADAS-2),[24] and observational studies (ROBINS-1).[25]
- Application of summary of findings, evidence profiles, and related GRADE methods has meant the characteristics of the intervention, characteristics of the participants, and risk of bias are more thoroughly and systematically documented.[26 27]
- Integration of other study designs containing evidence, such economic evaluation and qualitative research.[28]

There are other incremental improvements in a wide range of statistical and methodological areas, for example, in describing and taking into account cluster

Box 2: Examples of how different organisations decide on updating systematic reviews

Agency for Healthcare Research and Quality (US)

The AHRQ uses a needs based approach; updating systematic reviews depends on an assessment of several criteria:

1. Stakeholder impact
   - Interest from stakeholder partners (such as consumers, funders, guideline developers, clinical societies, James Lind Alliance)
   - Use and uptake (for example, frequency of citations and downloads)
   - Citation in scientific literature including clinical practice guidelines
2. Currency and need for update
   - New research is available
   - Review conclusions are probably dated
3. Update decision
   - Based on the above criteria, the decision is made to either update, archive, or continue surveillance.

Cochrane

Of over 50 Cochrane editorial teams, most but not all have some systems for updating, although this process can be informal and loosely applied. Most editorial teams draw on some or all of the following criteria:

1. Strategic importance
   - Is the topic a priority area (for example, in current debates or considered by guidelines groups)?
   - Is there important new information available?
2. Practicalities in organising the update that many groups take into account
   - Size of the task (size and quality of the review, and how many new studies or analyses are needed)
   - Availability and willingness of the author team
3. Impact of update
   - New research impact on findings and credibility
   - Consider whether new methods will improve review quality
4. Update decision
   - Priority to update, postpone update, class review as no longer requiring an update

ED_004886_00001537-00023

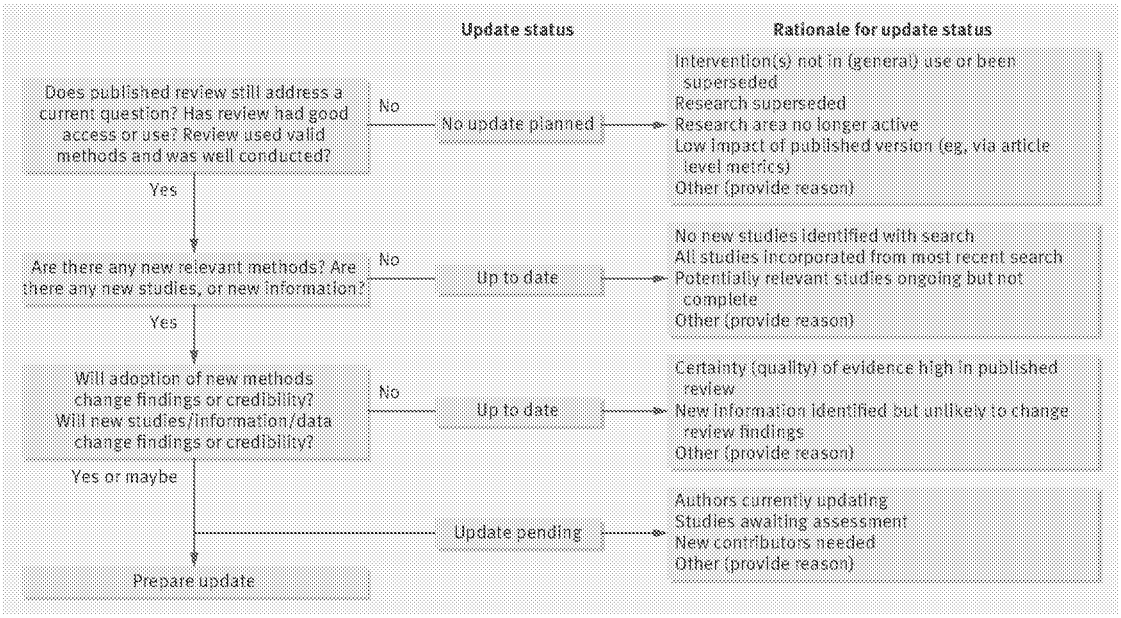


Fig 1 | Decision framework to assess systematic reviews for updating, with standard terms to report such decisions

randomised trials.[29] AMSTAR can assess the overall quality of a systematic review,[30] and the ROBIS tool can provide a more detailed assessment of the potential for bias.[31]

Are there any new studies or other information?

If an authoring or commissioning team wants to ensure that a particular review is up to date, there is a need for routine surveillance for new studies that are potentially relevant to the review, by searching and trial register inspection at regular intervals. This process has several approaches, including:

- Formal surveillance searching[32]
- Updating the full search strategies in the original review and running the searches
- Tracking studies in clinical trial and other registers
- Using literature appraisal services[33]
- Using a defined abbreviated search strategy for the update[34]
- Checking studies included in related systematic reviews.[35]

How often this surveillance is done, and which approaches to use, depend on the circumstances and the topic. Some topics move quickly, and the definition of "regular intervals" will vary according to the field and according to the state of evidence in the field. For example, early in the life of a new intervention, there might be a plethora of studies, and surveillance would be needed more frequently.

Box 3: Examples of a systematic review's currency

- The public is interested in vitamin C for preventing the common cold: the Cochrane review includes over 29 trials with either no or small effects, concluding good evidence of no important effects.[17] Assessment: still a current question for the public.
- Low osmolarity oral rehydration salt (ORS) solution versus standard solution for acute diarrhoea in children: the 2001 Cochrane review[18] led the World Health Organization to recommend ORS solution formula worldwide to follow the new ORS solution formula[19] and this has now been accepted globally. Assessment: no longer a current question.
- Routine prophylactic antibiotics with caesarean section: the Cochrane review reports clear evidence of maternal benefit from placebo controlled trials but no information on the effects on the baby.[20] Assessment: this is a current question.
- A systematic review published in the *Lancet* examined the effects of artemisinin based combination treatments compared with monotherapy for treating malaria and showed clear benefit.[21] Assessment: this established the treatment globally and is no longer a current question and no update is required.
- A Cochrane review of amalgam restorations for dental caries[22] is unlikely to be updated because the use of dental amalgam is declining, and the question is not seen as being important by many dental specialists. Assessment: no longer a current question.

Step 3: assess the effect of updating the review

Will the adoption of new methods change the findings or credibility?

Editors, referees, or experts in the topic area or methodologists can provide an informed view of whether a review can be substantially improved by application of current methodological expectations and new methods (fig 1). For example, a Cochrane review of iron supplementation in malaria concluded that there was "no significant difference between iron and placebo detected."[36] An update of the review included a GRADE assessment of the certainty of the evidence, and was able to conclude with a high degree of certainty that iron does not cause an excess of clinical malaria because the upper relative risk confidence intervals of harm was 1.0 with high certainty of evidence.[37]

Will the new studies, information, or data change the findings or credibility?

The assessment of new data contained in new studies and how these data might change the review is often used to determine whether an update should go ahead,

doi: 10.1136/bmj.i3507 | *BMJ* 2016;354:i3507 | the bmj

ED_004886_00001537-00024

Table 1 | Formal prediction tools: how potentially relevant new studies can affect review conclusions

Method	Description of approach	How it could be used	Advantages	Limitations	Validation
GRADE approach[38]	Considers whether the evidence certainty might change in the update (for example, because of lack of high certainty evidence, or because new evidence contradicts existing high certainty evidence). High certainty of evidence for critical outcomes could lower the priority for updating. Uncertainty in the review findings increases the need to include new studies[38]	Provides a benchmark by outcome to assess whether a new trial will improve the certainty of the evidence	Pragmatic. Many reviews already include GRADE	Requires GRADE to have been used in existing review or to complete an assessment according to GRADE	GRADE summary of findings tables or evidence profiles widely validated. Use of GRADE approach to prioritising updates requires further validation
Ottawa method[39-41]	A simple PubMed search (using the three largest and three most recent trials from the original review) to identify new studies. If new studies are found, then the method uses quantitative signals (eg, change in significance, effect size) to assess the likelihood that the new studies will change the review conclusion, thus triggering an update	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions; reasonable agreement with RAND method[39 42-44]
RAND method[45]	An abbreviated search of five major journals to identify new studies, and a search of the US Food and Drug Administration website and external expert judgment to determine the currency of the report findings	Practical routine surveillance tool	Easy to use	Will not detect all trials; judgment only based on changing conclusion	Approach validated for consistency of predicted and actual changes to conclusions, and compares well with the Ottawa method[39 43 44]
Statistical prediction tool[15]	A multicomponent decision tool to assess whether there might be any new studies for the update. If new studies are identified, a statistical prediction tool estimates the probability that this will change the review conclusion	Ranks multiple systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires commercial software	Approach validated internally[46]; requires further external validation
Value of information analysis[47 48]	Builds on the statistical prediction tool approach[15] comparing the expected health gain from new evidence with its expected cost. The gain is calculated in terms of a reduction in expected loss from reduced uncertainty and the cost is measured in days required to update the review.	Ranks selected systematic reviews in order of priority for updating	Uses quantitative approach	More complicated; requires some statistical knowledge	Approach validated internally; requires further external validation

Box 4: Examples of new information other than new trials being important

- The iconic Cochrane review of steroids in preterm labour was thought to provide evidence of benefit in infants, and this question no longer required new trials. However, a new large trial published in the *Lancet* in 2015 showed that in low and middle income countries, strategies to promote the uptake of neonatal steroids increased neonatal mortality and suspected maternal infection.[49] This information needs to somehow be incorporated into the review to maintain its credibility.
- A Cochrane review of community deworming in developing countries indicates that in recent studies, there is little or no effect.[50] The inclusion of a large trial of two million children confirmed that there was no effect on mortality. Although the incorporation of the trial in the review did not change the review's conclusions, the trial's absence would have affected the credibility of the review, so it was therefore updated.
- A new paper reporting long term follow-up data on anthracycline chemotherapy as part of cancer treatment was published. Although the effects from the outcomes remained essentially unchanged, apart from this longer follow-up, the paper also included information about the performance bias in the original trial, shifting the risk of bias for several outcomes from "unknown" to "high" in the Cochrane review.[51]

and the speed with which the update should be conducted. The appraisal of these new data can be carried out in different ways. Initially, methods focused on statistical approaches to predict an overturning of the current review findings in terms of the primary or desired outcome (table 1). Although this aspect is important, additional studies can add important information to a review, which is more than just changing the primary outcome to a more accurate and reliable estimate. Box 4 gives examples.

Reviews with a high level of certainty in the results (that is, when the GRADE assessment for the body of evidence is high) are less likely to change even with the addition of new studies, information, or data, by definition. GRADE can help guide priorities in whether to update, but it is still important to assess new studies that might meet the inclusion criteria. New studies can show unexpected effects (eg, attenuation of efficacy) or provide new information about the effects seen in different circumstances (eg, groups of patients or locations).

Other tools are specifically designed to help decision making in updating. For example, the Ottawa[39] and RAND[45] methods focus on identification of new evidence, the statistical predication tool[15] calculates the probability of new evidence changing the review conclusion, and the value of information analysis approach[52] calculates the expected health gain (table 1). As yet, there has been limited external validation of these tools to determine which approach would be most effective and when.

If potentially relevant studies are identified that have not previously been assessed for inclusion, authors or those managing the updating process need to assess whether including them might affect the conclusions of the review. They need to examine the weight and certainty of the new evidence to help determine whether an update is needed and how urgent that update is. The updating team can assess this informally by judging whether new studies or data are likely to substantively affect the review, for example, by altering the certainty in an existing comparison, or by generating new comparisons and analyses in the existing review.

New information can also include fresh follow-up data on existing included studies, or information on how the studies were carried out. These should be assessed in terms of whether they might change the review findings or improve its credibility (fig 1). Indeed, if any study has been retracted, it is important the authors assess the reasons for its retraction. In the case of data fabrication, the study needs to be removed

ED_004886_00001537-00025

from the analysis and this recorded. A decision needs to be made as to whether other studies by the same author should be removed from the review and other related reviews. An investigation should also be initiated following guidelines from the Committee on Publication Ethics (COPE). Additional published and unpublished data can become available from a wide range of sources—including study investigators, regulatory agencies and industry—and are important to consider.

Preparing for an update

Refresh background, objectives, inclusion criteria, and methods

Before including new studies in the review, authors need to revisit the background, objectives, inclusion criteria, and methods of the current review. In Cochrane, this is referred to as the protocol, and editors are part of this process. The update could range from simply endorsing the current question and inclusion criteria, through to full rewriting of the question, inclusion criteria and methods, and republishing the protocol. As a field progresses with larger and better quality trials rigorously testing the questions posed, it may be appropriate to exclude weaker study designs (such as quasi-randomised comparisons or very small trials) from the update (table 2). The PUGs panel recommended that a protocol refresh will require the authors to use the latest accepted methods of synthesis, even if this means repeating data extraction for all studies.

New authors and authorship

Updated systematic reviews are new publications with new citations. An authorship team publishing an update in a scientific or medical journal is likely to manage the new edition of a review in the same way as with any other publication, and follow the ICMJE authorship criteria.[56] If the previous author or author team steps down, then they should be acknowledged in the new version. However, some might perceive that their efforts in the first version warrant continued authorship, which may be valid. The management of authorship between versions can sometimes be complicated. At worst, it delays new authors completing an update and leads to long authorship lists of people from previous versions who probably do not meet ICMJE authorship criteria. One approach with updates including new authors is to have an opt-in policy for the existing authors: they can opt in to the new edition, provided that they make clear their contribution, and this is then agreed with the entire author team.

Although they are new publications, updates will generally include content from the published version. Changing licensing rights around systematic reviews to allow new authors of future updates to remix, tweak, or build on the contributions of the original authors of the published version (similar to the rights available via a Creative Commons licence; https://creativecommons.org) could be a more sustainable and simpler approach. This approach would allow systematic reviews to continue to evolve and build on the work of a range of authors over time, and for contributors to be given credit for contributions to this previous work.

Efficient searching

In performing an update, a search based on the search conducted for the original review is required. The updated search strategy will need to take into account changes in the review question or inclusion criteria, for example, and might be further adjusted based on knowledge of running the original search strategy. The search strategy for an update need not replicate the original search strategy, but could be refined, for example, based on an analysis of the yield of the original search. These new search approaches are currently undergoing formal empirical evaluation, but they may well provide much more efficient search strategies in the future. Some examples of these possible new methods for review updates are described in web appendix 2.

In reporting the search process for the update, investigators must ensure transparency for any previous versions and the current update, and use an adapted flow diagram based on PRISMA reporting (preferred reporting items for systematic reviews and meta-analyses).[57] The search processes and strategies for the update must be adequately reported such that they could be replicated.

Table 2 | Refresh background, objectives, inclusion criteria, and methods

Protocol section	Appraisal points
Background and research question	• Review and update background section, including supporting references to take account of any changes that may have occurred. This should include updating any new information and current policy debates on the topic. • Assess whether the current review question remains relevant to patients and practice.
Inclusion criteria	• Consider whether the existing PICO(s) remain(s) current, in the light of new knowledge. • Identify any new understanding of definition of patient populations. • Identify new interventions, or those that have been withdrawn, that are no longer in use. • Identify any changes in usual care standards. • Check for standardised core outcomes sets, such as those developed in collaboration with the core outcome measures in effectiveness trials (COMET) initiative (www.comet-initiative.org) or by guideline groups since the original review. • Check for any relevant patient reported outcomes to include subsequent to the original review. • Consider any new studies with less risk of bias that might warrant a stricter study design inclusion criteria (where the older version, when there was a dearth of evidence, included observational or quasi-randomised comparisons).
Methods	• Appraise and update the methods pending relevant methodological advancements or developments. For example, if (1) there are new tools for assessing the risk of bias of individual studies or appraising the quality of a body of evidence (eg, GRADE); or (2) new and efficient search approaches are feasible, such as a targeted approach to searching, taking into account the quality of the original search, and ensuring that the search for the update is of high quality. • Update or include a summary of findings table, which is recommended for all systematic reviews, because it improves the clarity, understanding, and interpretation of the findings of a systematic review, and rapidly reduces the amount of time readers require to find key information.[53-55]

doi: 10.1136/bmj.i3507 | BMJ 2016;354:i3507 | the bmj

ED_004886_00001537-00026

Peer review

Systematic reviews published for the first time in peer reviewed journals are by definition peer reviewed, but practice for updates remains variable, because an update might have few changes (such as an updated search but no new studies found and therefore included) or many changes (such as revise methods and inclusion of several new studies leading to revised conclusions). Therefore, and to use peer reviewers' time most effectively, editors need to consider when to peer review an update and the type of peer reviewer most useful for a particular update (for example, topic specialist, methodologist). The decision to use peer review, and the number and expertise of the peer reviewers could depend on the nature of the update and the extent of any changes to the systematic review as part of an editor assessment. A change in the date of the search only (where no new studies were identified) would not require peer review (except, arguably, peer review of the search), but the addition of studies that lead to a change in conclusions or significant changes to the methods would require peer review. The nature of the peer review could be described within the published article.

Reporting changes

Authors should provide a clear description of the changes in approach or methods between different editions of a review. Also, authors need to report the differences in findings between the original and updated edition to help users decide how to use the new edition. The approach or format used to present the differences in findings might vary with the target user group.[58] Publishers need to ensure that all previous versions of the review remain publically accessible.

Updates can range from small adjustments to reviews being completely rewritten, and the PUGs panel spent some time debating whether the term "new edition" would be a better description than "update." However, the word "update" is now in common parlance and changing the term, the panel judged, could cause confusion. However, the debate does illustrate that an update could represent a review that asks a similar question but has been completely revised.

Technology and innovation

The updating of systematic review is generally done manually and is time consuming. There are opportunities to make better use of technology to streamline the updating process and improve efficiency (table 3). Some of these tools already exist and are in development or in early use, and some are commercially available or freely available. The AHRQ's evidence based practice centre team has recently published tools for searching and screening, and will provide an assessment of the use, reliability, and availability of these tools.[63]

Other developments, such as targeted updates that are performed rapidly and focus on updating only key components of a review, could provide different approaches to updating in the future and are being piloted and evaluated.[64] With implementation of these various innovations, the longer term goal is for "living" systematic reviews, which identify and incorporate information rapidly as it evolves over time.[60]

Concluding remarks

Updating systematic reviews, rather than addressing the same question with a fresh protocol, is generally

Table 3 | Technological innovations to improve the efficiency of updating systematic reviews

Innovation	Description	Application	Examples of software and projects,* and current status
Integrated software	Integration of applying inclusion criteria, review management systems, statistical packages, and GRADE	To facilitate greater efficiencies in review production, including their updates	Covidence (www.covidence.org): free/pay† EPPI reviewer (http://eppi.ioe.ac.uk/cms/Default.aspx): pay DistillerSR (https://distillercer.com/products/distillersr-systematic-review-software/): pay Cochrane Review Manager (http://tech.cochrane.org/revman): free/pay† GRADEpro GDT (http://gradepro.org/): free Rayyan (http://rayyan.qcri.org/): free
Systematic review data repositories	Repositories store information from review (eg, data abstraction forms and the evidence tables)	Improve updating efficiency for new or existing teams as the data abstraction forms, evidence tables, and populating data from the original review are available	Agency of Health Care Research and Quality systematic review data repository (http://srdr.ahrq.gov): operational GRADE database of evidence profiles and evidence to decision frameworks (http://dbep.gradepro.org/): operational
Semi-automation	Machine learning techniques to use alongside human efforts	Finding studies and extracting data could benefit from semi-automation creating time efficiencies[59]	RobotReviewer (http://vortext.systems/robotreviewer): free Cochrane project transform—identification of studies (http://community.cochrane.org/tools/project-coordination-and-support/transform)
Crowdsourcing	Use of volunteers to assist systematic review authors with discrete tasks	Individuals from the "crowd" assist with tasks (identifying and screening studies, translating articles, data extraction) to help in new review production and updates[60-62]	Cochrane Project Transform—crowdsourcing (link as above)
Publication linkage	Ability to link trial registration, trial publications, and reviews citing them will help transparency	This initiative could help identify studies for systematic reviews and could also show the relation between systematic review updates	A cross publisher initiative, CrossRef, is coordinating a threaded publications/linked clinical trial reports initiative to link a clinical trial report (with a trial registration number) report and derivative publications, including reviews (www.crossref.org): operational
Data linkage	Increase links between data, existing software, and reviews	To improve identification and reuse of data for review production and dissemination	Cochrane (http://linkeddata.cochrane.org/): proof of principle example at production phase, but mostly linkage projects at exploratory phase

*Further information can be located on the SR Toolbox site (http://systematicreviewtools.com/).
†Free to Cochrane contributors; other users pay.

ED_004886_00001537-00027

more efficient and allows incremental improvement over time. Mechanical rules appear unworkable, but there is no clear unified approach on when to update, and how implement this. This PUGs panel of authors, editors, statisticians, information specialists, other methodologists, and guideline developers brought together current thinking and experience in this area to provide guidance.

Decisions about whether and when to update a systematic review are judgments made at a point in time. They depend on the currency of the question asked, the need for updating to maintain credibility, the availability of new evidence, and whether new research or new methods will affect the findings.

Whether the review uses current methodological standards is important in deciding if the update will influence the review findings, quality, reliability, or credibility sufficiently to justify the effort in updating it. Those updating systematic reviews to author clinical practice guidelines might consider the influence of new study results in potentially overturning the conclusions of an existing review. Yet, even in cases where new study findings do not change the primary outcome measure, new studies can carry important information about subgroup effects, duration of treatment effects, and other relevant clinical information, enhancing the currency and breadth of review results.

An update requires appraisal and revision of the background, question, inclusion criteria, and methods of the existing review and the existing certainty in the evidence. In particular, methods might need to be updated, and search strategies reconsidered. Authors of updates need to consider inputs to the current edition, and follow ICMJE criteria regarding authorship.[56]

The PUGs panel proposed a decision framework (fig 1), with terms and categories for reporting the decisions made for updating procedures for adoption by Cochrane and other stakeholders. This framework includes journals publishing systematic review updates and independent authors considering updates of existing published reviews. The panel developed a checklist to help judgements about when and how to update.

The current emphasis of authors, guideline developers, Cochrane, and consequently this guidance has been on effects reviews. The checklists and guidance here still applies to other types of systematic reviews, such as those on diagnostic test accuracy, and this guidance will need adapting. Accumulative experience and methods development in reviews other than those of effects are likely to help refine guidance in the future.

This guidance could help groups identify and prioritise reviews for updating and hence use their finite resources to greatest effect. Software innovation and new management systems are being developed and in early use to help streamline review updates in the coming years.

AUTHOR AFFILIATIONS

[1]Cochrane Infectious Diseases Group, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK
[2]Oxford Clinical Trials Research Unit, University of Oxford, Oxford, UK
[3]Cochrane Editorial Unit, Cochrane Central Executive, London, UK
[4]Department of Clinical Epidemiology and Biostatistics and Department of Medicine, McMaster University, Hamilton, ON, Canada
[5]Cochrane GRADEing Methods Group, Ottawa, ON, Canada
[6]Department of Internal Medicine, American University of Beirut, Beirut, Lebanon
[7]Department of Mathematics and Statistics, McMaster University
[8]Evidence-based Practice Center Program, Agency for Healthcare and Research Quality, Rockville, MD, USA
[9]Centre for Reviews and Dissemination, University of York, York, UK
[10]Cochrane Upper Gastrointestinal and Pancreatic Diseases Group, Hamilton, ON, Canada
[11]Lefebvre Associates, Oxford, UK
[12]Kaiser Permanente National Guideline Program, Portland, OR, USA
[13]Iberoamerican Cochrane Centre, Barcelona, Spain
[14]Cochrane Informatics and Knowledge Management, Cochrane Central Executive, Freiburg, Germany
[15]Plymouth University Peninsula School of Dentistry, Plymouth, UK
[16]Department of Clinical Policy, American College of Physicians,Philadelphia, PA, USA
[17]Guidelines International Network, Pitlochry, UK
[18]Children's Hospital of Eastern Ontario, Ottawa, ON, Canada
[19]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[20]Biostatistics Unit, Centre for Evaluation, McMaster University, Hamilton, ON, Canada
[21]Centre for Statistics in Medicine, University of Oxford, Oxford, UK
[22]University of Ottawa, Ottawa, ON, Canada
[23]Cochrane Airways Group, Population Health Research Institute, St George's, University of London, London, UK
[24]Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, UK

Contributors: HJS initiated the workshop. JC, SH, PG, HM, and HJS organised the materials and the agenda. SH wrote up the proceedings. PG wrote the paper from the proceedings and coordinated the development of the final guidance; JC, SH, HM, and HJS were active in the finalising of the guidance. All PUGs authors contributed to three rounds of manuscript revision.

Funding: Attendance at this meeting, for those attendees not directly employed by Cochrane, was not funded by Cochrane beyond the reimbursement of out of pocket expenses for those attendees for whom this was appropriate. Expenses were not reimbursed for US federal government attendees, in line with US government policy. Statements in the manuscript should not be construed as endorsement by the US Agency for Healthcare Research and Quality or the US Department of Health and Human Services.

Competing interests: All participants have a direct or indirect interest in systematic reviews and updating as part of their job or academic career. Most participants contribute to Cochrane, whose mission includes a commitment to the updating of its systematic review portfolio. JC, HM, RM, CM, KS-W, and MT are, or were at that time, employed by the Cochrane Central Executive.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 3.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/3.0/.

1 Shekelle PG, Ortiz E, Rhodes S, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated?*JAMA* 2001;286:1461-7. doi:10.1001/jama.286.12.1461.
2 Claxton K, Cohen JT, Neumann PJ. When is evidence sufficient?*Health Aff (Millwood)* 2005;24:93-101. doi:10.1377/hlthaff.24.1.93.
3 Fenwick E, Claxton K, Sculpher M, et al. *Improving the efficiency and relevance of health technology assessment: the role of decision analytic modelling. Paper 179.*Centre for Health Economics, University of York, 2000.
4 Sculpher M, Claxton K. Establishing the cost-effectiveness of new pharmaceuticals under conditions of uncertainty—when is there sufficient evidence?*Value Health* 2005;8:433-46. doi:10.1111/j.1524-4733.2005.00033.x.

ED_004886_00001537-00028

5 Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *J Health Serv Res Policy* 1997;2:26-30.
6 Wilson E, Abrams K. From evidence based economics to economics based evidence: using systematic review to inform the design of future research. In: Shemilt I, Mugford M, Vale L, et al. eds. *Evidence based economics*.Blackwell Publishing, 2010doi:10.1002/9781444320398.ch12.
7 Chalmers I, Enkin M, Keirse MJ. Preparing and updating systematic reviews of randomized controlled trials of health care. *Milbank Q* 1993;71:411-37. doi:10.2307/3350409.
8 Higgins J, Green S, Scholten R. Chapter 3. *Maintaining reviews: updates, amendments and feedback: Version 5.1.0 (updated March 2011)*.Cochrane Collaboration, 2011.
9 Cochrane. Editorial and publishing policy resource. http://community.cochrane.org/editorial-and-publishing-policy-resource. 2016.
10 Moher D, Tsertsvadze A. Systematic reviews: when is an update an update?*Lancet* 2006;367:881-3. doi:10.1016/S0140-6736(06)68358-X.
11 Prasad K, Singh MB, Ryan H. Corticosteroids for managing tuberculous meningitis. *Cochrane Database Syst Rev* 2016;4:CD002244.
12 Zani B, Gathu M, Donegan S, Olliaro PL, Sinclair D. Dihydroartemisinin-piperaquine for treating uncomplicated *Plasmodium falciparum* malaria. *Cochrane Database Syst Rev* 2014;1:CD010927.
13 Adams SP, Tsang M, Wright JM. Lipid lowering efficacy of atorvastatin. *Cochrane Database Syst Rev* 2012;12:CD008226.
14 Higgins J. Convincing evidence from controlled and uncontrolled studies on the lipid-lowering effect of a statin. *Cochrane Database Syst Rev* 2012;12:ED000049.
15 Takwoingi Y, Hopewell S, Tovey D, Sutton AJ. A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ* 2013;347:f7191. doi:10.1136/bmj.f7191.
16 MacLehose H, Hilton J, Tovey D, et al. The Cochrane Library: revolution or evolution? Shaping the future of Cochrane content. Background paper for The Cochrane Collaboration's Strategic Session Paris, France, 18 April 2012. http://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/2012-CC-strategic-session_full-report.pdf.
17 Hemilä H, Chalker E. Vitamin C for preventing and treating the common cold. *Cochrane Database Syst Rev* 2013;(1):CD000980.
18 Hahn S, Kim S, Garner P. Reduced osmolarity oral rehydration solution for treating dehydration caused by acute diarrhoea in children. *Cochrane Database Syst Rev* 2002;(1):CD002847.
19 World Health Organization (WHO). Reduced osmolarity oral rehydration salts (ORS) formulation. A report from a meeting of Experts jointly organized by UNICEF and WHO. New York: Child and Adolescent Health and Development, 18 July 2001 http://apps.who.int/iris/bitstream/10665/67322/1/WHO_FCH_CAH_01.22.pdf.
20 Smaill FM, Grivell RM. Antibiotic prophylaxis versus no prophylaxis for preventing infection after cesarean section. *Cochrane Database Syst Rev* 2014;(10):CD007482.
21 Adjuik M, Babiker A, Garner P, Olliaro P, Taylor W, White N. International Artemisinin Study Group. Artesunate combinations for treatment of malaria: meta-analysis. *Lancet* 2004;363:9-17. doi:10.1016/S0140-6736(03)15162-8.
22 Agnihotry A, Fedorowicz Z, Nasser M. Adhesively bonded versus non-bonded amalgam restorations for dental caries. *Cochrane Database Syst Rev* 2016;3:CD007517.
23 Higgins JP, Altman DG, Gøtzsche PC, et al. Cochrane Bias Methods Group Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928. doi:10.1136/bmj.d5928.
24 Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36. doi:10.7326/0003-4819-155-8-201110180-00009.
25 Sterne JAC, Higgins JPT, Reeves BC; on behalf of the development group for ROBINS-I. A tool for assessing risk of bias in non-randomized studies of interventions, version 7. March 2016. www.riskofbias.info.
26 Guyatt GH, Oxman AD, Vist GE, et al. GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-6. doi:10.1136/bmj.39489.470347.AD.
27 Schünemann HJ. Interpreting GRADE's levels of certainty or quality of the evidence: GRADE for statisticians, considering review information size or less emphasis on imprecision?*J Clin Epidemiol* 2016;75:6-15. doi:10.1016/j.jclinepi.2016.03.018.
28 Gough D. Qualitative and mixed methods in systematic reviews. *Syst Rev* 2015;4:181. doi:10.1186/s13643-015-0151-y.
29 Richardson M, Garner P, Donegan S. Cluster randomised trials in Cochrane reviews: evaluation of methodological and reporting practice. *PLoS One* 2016;11:e0151818. doi:10.1371/journal.pone.0151818.
30 Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol* 2007;7:10. doi:10.1186/1471-2288-7-10.
31 Whiting P, Savović J, Higgins JP, et al. ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005.
32 Sampson M, Shojania KG, McGowan J, et al. Surveillance search techniques identified the need to update systematic reviews. *J Clin Epidemiol* 2008;61:755-62. doi:10.1016/j.jclinepi.2007.10.003.
33 Hemens BJ, Haynes RB. McMaster Premium LiteratUre Service (PLUS) performed well for identifying new studies for updated Cochrane reviews. *J Clin Epidemiol* 2012;65:62-72.e1. doi:10.1016/j.jclinepi.2011.02.010.
34 Sagliocca L, De Masi S, Ferrigno L, Mele A, Traversa G. A pragmatic strategy for the review of clinical evidence. *J Eval Clin Pract* 2013;19:689-96. doi:10.1111/jep.12020.
35 Rada G, Peña J, Capurro D, et al. How to create a matrix of evidence in Epistemonikos. Abstracts of the 22nd Cochrane Colloquium; Evidence-informed public health: opportunities and challenges; Hyderabad, India. *Cochrane Database Syst Rev* 2014;suppl 1:132.
36 Okebe JU, Yahav D, Shbita R, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2011;(10):CD006589.
37 Neuberger A, Okebe J, Yahav D, Paul M. Oral iron supplements for children in malaria-endemic areas. *Cochrane Database Syst Rev* 2016;2:CD006589.
38 Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401-6. doi:10.1016/j.jclinepi.2010.07.015.
39 Chung M, Newberry SJ, Ansari MT, et al. Two methods provide similar signals for the need to update systematic reviews. *J Clin Epidemiol* 2012;65:660-8. doi:10.1016/j.jclinepi.2011.12.004.
40 Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med* 2007;147:224-33. doi:10.7326/0003-4819-147-4-200708210-00179.
41 Shojania K, Sampson M, Ansari M, et al. *Updating systematic reviews; AHRQ technical reviews; report no 07-0087*.Agency for Healthcare Research and Quality, 2007.
42 Pattanittum P, Laopaiboon M, Moher D, Lumbiganon P, Ngamjarus C. A comparison of statistical methods for identifying out-of-date systematic reviews. *PLoS One* 2012;7:e48894. doi:10.1371/journal.pone.0048894.
43 Shekelle PG, Motala A, Johnsen B, Newberry SJ. Assessment of a method to detect signals for updating systematic reviews. *Syst Rev* 2014;3:13. doi:10.1186/2046-4053-3-13.
44 Shekelle PG, Newberry SJ, Wu H, et al. *Identifying signals for updating systematic reviews: a comparison of two methods; report no 11-EHC042-EF*.Agency for Healthcare Research and Quality, 2011.
45 Shekelle P, Newberry S, Maglione M, et al. *Assessment of the need to update comparative effectiveness reviews: report of an initial rapid program assessment (2005-2009)*.Agency for Healthcare Research and Quality, 2009.
46 Tovey D, Marshall R, Bazian, Hopewell S, Rader T. Fit for purpose: centralised updating support for high-priority Cochrane Reviews; National Institute for Health Research Cochrane-National Health Service Engagement Award Scheme, July 2011. https://editorial-unit.cochrane.org/sites/editorial-unit.cochrane.org/files/uploads/10_4000_01%20Fit%20for%20purpose%20-%20centralised%20updating%20support%20for%20high%20priority%20Cochrane%20Reviews%20FINAL%20REPORT.pdf.
47 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;18:341-64. doi:10.1016/S0167-6296(98)00039-3.
48 Wilson EC. A practical guide to value of information analysis. *Pharmacoeconomics* 2015;33:105-21. doi:10.1007/s40273-014-0219-x.
49 Althabe F, Belizán JM, McClure EM, et al. A population-based, multifaceted strategy to implement antenatal corticosteroid treatment versus standard care for the reduction of neonatal mortality due to preterm birth in low-income and middle-income countries: the ACT cluster-randomised trial. *Lancet* 2015;385:629-39. doi:10.1016/S0140-6736(14)61651-2.
50 Taylor-Robinson D, Maayan N, Soares-Weiser K, et al. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin and school performance. *Cochrane Database Syst Rev* 2012;(11):CD000371.
51 van Dalen EC, van der Pal HJ, Kremer LC. Different dosage schedules for reducing cardiotoxicity in people with cancer receiving anthracycline chemotherapy. *Cochrane Database Syst Rev* 2016;3:CD005008.
52 Wilson E. on behalf of the Cochrane Priority Setting and Campbell & Cochrane Economics Methods Groups. Which study when? Proof of concept of a proposed automated tool to help decision which reviews to update first. *Cochrane Database Syst Rev* 2014;suppl 2:29-31.
53 Rosenbaum SE, Glenton C, Nylund HK, Oxman AD. User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews. *J Clin Epidemiol* 2010;63:607-19. doi:10.1016/j.jclinepi.2009.12.013.
54 Rosenbaum SE, Glenton C, Oxman AD. Summary-of-findings tables in Cochrane reviews improved understanding and rapid retrieval of key information. *J Clin Epidemiol* 2010;63:620-6. doi:10.1016/j.jclinepi.2009.12.014.

ED_004886_00001537-00029

55 Vandvik PO, Santesso N, Akl EA, et al. Formatting modifications in GRADE evidence profiles improved guideline panelists comprehension and accessibility to information. A randomized trial. *J Clin Epidemiol* 2012;65:748-55. doi:10.1016/j.jclinepi.2011.11.013.

56 International Committee of Medical Journal Editors (ICMJE). Defining the role of authors and contributors. 2016. www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html.

57 Stovold E, Beecher D, Foxlee R, Noel-Storr A. Study flow diagrams in Cochrane systematic review updates: an adapted PRISMA flow diagram. *Syst Rev* 2014;3:54. doi:10.1186/2046-4053-3-54.

58 Newberry SJ, Shekelle PG, Vaiana M, et al. *Reporting the findings of updated systematic reviews of comparative effectiveness: how do users want to view new information? report no 13-EHC093-EF* Agency for Healthcare Research and Quality, 2013.

59 Marshall IJ, Kuiper J, Wallace BC. Automating risk of bias assessment for clinical trials BCB'14. Proceedings of the 5th ACM conference on Bioinformatics, computational biology, and health informatics. 2014:88-95. http://thirdworld.nl/automating-risk-of-bias-assessment-for-clinical-trials.

60 Elliott JH, Turner T, Clavisi O, et al. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014;11:e1001603. doi:10.1371/journal.pmed.1001603.

61 Elliott J, Sim I, Thomas J, et al. #CochraneTech: technology and the future of systematic reviews. *Cochrane Database Syst Rev* 2014;(9):ED000091.

62 Cochrane. Project transform: the Cochrane Collaboration. 2016. http://community.cochrane.org/tools/project-coordination-and-support/transform.

63 Paynter R, Bañez L, Berlinerm E, et al. EPC methods: an exploration of the use of text-mining software in systematic reviews. Research white paper. AHRQ publication 16-EHC023-EF. Agency for Healthcare Research and Quality, April 2016. https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2214.

64 Soares-Weiser K, Marshall R, Bergman H, et al. Updating Cochrane Reviews: results of the first pilot of a focused update. *Cochrane Database Syst Rev* 2014;suppl 1:31-3.

© BMJ Publishing Group Ltd 2016

Web appendix: Supplementary material

No commercial reuse: See rights and reprints http://www.bmj.com/permissions Subscribe: http://www.bmj.com/subscribe