Science of the Total Environment 466–467 (2014) 315–325

Contents lists available at ScienceDirect

Science of the Total Environment

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

The Matthew Effect and widely prescribed pharmaceuticals lacking environmental monitoring: Case study of an exposure-assessment vulnerability

Christian G. Daughton *

Environmental Sciences Division, National Exposure Research Laboratory, U.S. Environmental Protection Agency, 944 East Harmon Avenue, Las Vegas, NV 89119, USA

HIGHLIGHTS

- Targeting of chemicals for environmental monitoring is often biased by prior data.
- The Matthew Effect may act to exclude certain chemicals as targets for monitoring.
- Chemicals not previously targeted for monitoring can become perpetually ignored.
- Exposure assessment may be vulnerable to bias created by the Matthew Effect.
- Protocols to counter the Matthew Effect may improve environmental decision making.

ARTICLE INFO

Article history:
Received 4 June 2013
Received in revised form 26 June 2013
Accepted 26 June 2013
Available online 1 August 2013

Editor: Damia Barcelo

Keywords:
Risk assessment

ABSTRACT

Assessing ambient exposure to chemical stressors often begins with time-consuming and costly monitoring studies to establish environmental occurrence. Both human and ecological toxicology are currently challenged by the unknowns surrounding low-dose exposure/effects, compounded by the reality that exposure undoubtedly involves mixtures of multiple stressors whose identities and levels can vary over time. Long absent from the assessment process, however, is whether the full scope of the identities of the stressors is sufficiently known. The Matthew Effect (a psychosocial phenomenon sometimes informally called the "bandwagon effect" or "iceberg effect," among others) may adversely bias or corrupt the exposure assessment process. The Matthew Effect is evidenced by decisions that base the selection of stressors to target in environmental monitoring surveys on whether they have been identified in prior studies, rather than considering the pos-

research is explored for the first time in a comprehensive case study that examines the preponderance of "absence of data" (in contrast to positive data and "data of absence") for the environmental occurrence of a very large class of potential chemical stressors associated with ubiquitous consumer use — active pharmaceutical ingredients (APIs). Comprehensive examination of the published data for an array of several hundred of the most frequently used drugs for whether their APIs are environmental contaminants provides a prototype example to catalyze discussion among the many disciplines involved with assessing risk. The findings could help guide the selection of those APIs that might merit targeting for environmental monitoring (based on the absence of data for environmental occurrence) as well as the prescribing of those medications that might have minimal environmental impact (based on data of absence for environmental occurrence).

1. Introduction/background

Historically, the spectrum of chemical stressors considered in risk assessments has been narrowly restricted to regulated priority and legacy pollutants and associated conventional chemicals, such as high-volume commercial products or those unintentionally produced as ubiquitous by-products from industrial processes (Daughton and Ternes, 1999). These numbers are comparatively few, however, among the tens of thousands in widespread commercial use, without even considering

Abbreviations: API, active pharmaceutical ingredient; EPA, U.S. Environmental Protection Agency; IRIS, EPA's Integrated Risk Information System; MEOC, Matthew Effect Orphaned Chemical; PEC, Predicted Environmental Concentration.

* Tel.: +1 702 798 2207; fax: +1 702 798 2142.
E-mail address: daughton.christian@epa.gov.

0048-9697
http://dx.doi.org/10.1016/j.scitotenv.2013.06.111

the thousands of unique active ingredients used in pharmaceutical preparations — the APIs.

This tendency of comparatively few, select chemicals to occupy the attention of the many disciplines involved in risk assessment has been noted over the last decade, even in light of the more expansive perspective on environmental contamination afforded by the so-called "emerging" contaminants (such as APIs). A disparity in the scope of data from environmental monitoring continues to grow as a result of the comparatively few chemical stressors that are targeted. This phenomenon has been deemed a manifestation of a self-fulfilling selection bias referred to as the "Matthew Effect" — where the prominence of those few chemicals targeted for investigation is dictated largely by the attention devoted to them in the past. The Matthew Effect as a psychosocial phenomenon was first articulated by Robert Merton in his well-known work in 1968 (Merton, 1968) and later used by Grandjean et al. (2011) to explain the biased path followed by many of the incremental and repetitive findings of environmental science.

The psychology involved with the conduct of science (specifically the Matthew Effect) is largely unexplored as a factor that drives the direction of science while creating unrecognized (stealth) bias and uncertainty. At least one aspect of the Matthew Effect is a cognitive bias called "availability bias" (or "availability heuristic"), which can elicit an overestimation of risk or probability associated with an event the more easily and vividly the event (or a similar event) can be recalled. Events that are more easily recalled are more susceptible to being viewed as risky or probable. This in turn further impairs the already reduced focus on those events that are not easily recalled. In this manner, attention tends to become devoted to events (in this case, chemicals occurring as environmental contaminants) that are more familiar, while the majority becomes relegated to obscurity.

Although some of the potential ramifications of the Matthew Effect in environmental science have been discussed by Grandjean et al. (2011), evidence of it playing a possible role in introducing bias has never been actively sought. After all, establishing an absence of published data for a particular subject is clearly an onerous task demanding rigorous and time-consuming examination of as much of the published literature as possible. Trying to establish that something has not been reported is usually perceived as a thankless endeavor and may explain why the Matthew Effect (if indeed an active phenomenon) could escape notice. Perhaps the only published example that the Matthew Effect may be at play in environmental science comes from a 2013 study that examined the potential impact of medication prescribing practices on environmental contamination by APIs (Daughton and Ruhoy, 2013). Data were presented (Table S1, therein) showing a select group of APIs that are prescribed frequently and whether evidence exists that they also occur in the environment as contaminants. Among the 53 frequently prescribed APIs subject of that evaluation, minimal evidence existed in the published literature for whether roughly a dozen had ever been targets of environmental monitoring (an absence of data as opposed to data of absence or "evidence of absence"). This absence of environmental occurrence data for 22% of a small sampling of 53 commonly used medications indicated the possibility of a substantially greater incidence of the absence of data for the much larger universe of the thousands of distinct APIs in use today. Also note that even the limited data published for the remainder of these APIs may have little relevance to indicating overall environmental presence, as many of these data do not result from formal monitoring surveys but rather are generated from a select few, convenient real-world samples analyzed in the course of developing and testing an analytical method for chemical analysis.

1.1. The Matthew Effect — contributor to uncertainty in assessing risk

Given the unknown numbers of yet-to-be-identified xenobiotics that may play roles in the totality of biological exposure, it is critical that the select few that are targeted for environmental or biological monitoring serve as sufficient proxies for assessing risk. With the increasing sophistication of analytical chemistry, the "iceberg" scenario has become more evident (Daughton, in press) — where escalating numbers of unique chemicals may be present at ever-lower, and often immeasurable, concentrations. Emerging is evidence that collections of chemicals remaining unidentified (e.g., see: Tang et al., 2013) may hold the predominant share of overall total biological stress in some exposure scenarios (e.g., Escher et al., in press). The implications of the iceberg scenario can be exacerbated by the Matthew Effect — as the risk posed by chemicals currently not measured is effectively shielded from examination or evaluation.

The critical importance of exposure assessment and its role as the weak link in both ecological and human health risk assessment is made clear in the European Commission's 2012 report "Addressing the New Challenges for Risk Assessment" (SCENIHR et al., 2012 [8 October]). Notably, however, this report (like all prior evaluations of the risk assessment paradigm) perpetuates an extremely limited and biased view of the chemical space occupied by chemical stressors. The report does not entertain the question as to whether the universe of stressors is sufficiently known, nor does it recognize the potential for bias and data disparity by ignoring those chemical stressors lacking sufficient environmental occurrence data. After all, these stressors represent that portion of potential exposure which is stealth or masked — exposure that is unknown but not recognized as such.

This same oversight is evident in the recent US GAO (2013) report and the Institute of Medicine (IOM) report "Environmental Decisions in the Face of Uncertainty" (IOM, 2013), where the IOM was requested by the U.S. Environmental Protection Agency (EPA) to provide guidance on managing risk in the face of uncertainty. While the possibility of bias and uncertainly introduced by the Matthew Effect was not alluded to in the IOM report, an argument could be made that the Matthew Effect should be addressed on the basis of two of the questions posed in the charge to the IOM:

(1) "How does uncertainty influence risk management under different public health policy scenarios?"
(2) "Are there other ways in which the EPA can benefit from quantitative characterization of uncertainty?"

Two of the recommendations made by the IOM report that were elicited by these questions were:

(1) "To better inform the public and decision makers, EPA decision documents and other communications to the public should systematically … include information on what uncertainties in the health risk assessment are present…"
(2) "EPA should develop methods to systematically describe and account for uncertainties in decision-relevant factors…"

A methodology that facilitates a formal evaluation of whether the Matthew Effect might be actively biasing a risk assessment could at least in part address these two recommendations.

1.2. Goals and objectives

Presented here is an approach for assessing the potential importance of those chemical stressors that are essentially ignored because of selective focus on the few stressors that repeatedly attract attention due to the Matthew Effect. One objective is to examine a potential vulnerability in assessing exposure: the failure to evaluate chemicals orphaned by the Matthew Effect and the absence of data (absence of evidence). Here, the stressors that are orphaned by the Matthew Effect are termed "Matthew Effect Orphaned Chemicals" (MEOCs). A case study is used to: (i) define the scope of MEOCs that might exist for a diverse class of potential chemical stressors (using a case study encompassing the environmental contaminants originating from pharmaceutical usage — the APIs), (ii) develop a strategy for

avoiding the vulnerabilities stemming from MEOCs in assessing exposure, and (iii) raise awareness of the Matthew Effect and catalyze discussion on identifying and overcoming the Matthew Effect among those involved with all aspects of ecological and human health risk assessment. With a less-biased approach to selecting those stressors that should be considered for environmental monitoring, the integrity and credibility of exposure assessments could be improved and uncertainty could be reduced. The ultimate objective is to steer attention by researchers and regulators toward identifying those classes or individual chemicals that are currently ignored but which may play important roles in biological exposure — striving to achieve more holistic exposure assessments.

A case study is used to demonstrate the potential incidence of MEOCs among environmental contaminants. Its focus is on APIs derived from routine pharmaceutical usage. Examinations of API environmental occurrence data rarely attempt to systematically establish the absence of data. APIs were selected as a test case for examining the potential prevalence of the Matthew Effect for the following reasons:

(1) a large published literature exists on the occurrence and monitoring of APIs in a wide variety of environmental compartments and matrices;
(2) ready availability of a comprehensive bibliographic database of documents relevant to the many aspects of APIs as environmental contaminants (this database greatly facilitates full-text Boolean searches and maximizes coverage of both the published, in-press, and gray literature, including over 14,000 journal articles, book chapters, reports, and dissertations) (Daughton and Scuderi, 2013);
(3) data on relative usage rates of APIs is readily available (such as the top 100 or 200 most frequently prescribed drugs); this allows the targeting of the most widely used APIs – among the roughly 9000 now known to be in commercial use (Huang et al., 2011) – which therefore should be among those with high probabilities of gaining entry to the environment; also see the discussion on the number of distinct APIs in use today in Daughton (in press);
(4) APIs are exempt from the EU's REACH program (Holzer, 2010), which may increase the odds that environmental occurrence data are absent (because of a lack of incentive to obtain these data); and
(5) ongoing and repetitive investment of research resources may be needlessly wasted in a range of fate and toxicity studies for APIs that have not yet been targeted for environmental monitoring but which have a low probability of occurring in the environment (examples can be seen in Table 3, which is introduced and explained in Section 4). For these APIs, low priority might be given for fate and toxicity studies.

1.3. Liabilities and vulnerabilities associated with the Matthew Effect

Historically, the preferential reporting of positive versus negative (non-positive) data has troubled a number of disciplines — one in particular being risk assessment (Buffler, 1989). The concern has been whether publications sometimes omit the documented absence of adverse health effects as a result of a bias, whether resulting simply because such results are perceived as "uninteresting," suspect, or fail to support a hypothesis. Even though the recognition of the value of negative results has grown (and they figure prominently in debates surrounding subtle, low-dose effects), the two types of data provide only two perspectives on what is knowable — essentially yes–no answers to whatever questions are under investigation. A third perspective is equally important and that is whether the correct questions are being posed or whether a sufficiently comprehensive suite of variables probed. This perspective can remain hidden when bias concentrates the focus on previously demonstrated interesting or widely recognized findings.

Assessing the risks posed by chemicals as environmental contaminants requires first establishing parameters surrounding their presence in the environment (including locations, matrices, levels, and spatiotemporal distributions and fluctuations). For certain chemicals, however, these data can be completely lacking. The extent and magnitude of this deficiency can be unknown and, at the time, not recognized as such (so-called "unknown unknowns"). Worse is that no mechanism is in place to rectify this vulnerability.

One of a number of specific examples that might exhibit bias from the Matthew Effect involves EPA's Integrated Risk Information System, IRIS (US EPA, 2013b), which is generally recognized as a definitive, authoritative technical resource for assessing health effects, albeit for a limited number of chemicals used in – or resulting from – commerce; another example is the EPA's Chemical Data Reporting (CDR) database (US EPA, 2013a).

An unintended consequence of this recognition is that the absence from IRIS of a particular chemical can be misconstrued as evidence that it may not play a meaningful role as a stressor in exposure (e.g., see: Gray and Cohen, 2012). Such erroneous inferences result when the absence of data is conflated with evidence of absence. This is an example of "availability bias." Indeed, this is one of the tactics used in the so-called "Four Dog Defense" — where "my dog has the ability to bite, but it would never bite you" (Sass and Rosenberg, 2011 [October]). Or in other words, a chemical might pose a hazard, but concern is not warranted if the chemical is not present in the environment (a false assertion when based on an absence of data rather than on sufficient data of absence). In reality, when a chemical is not accommodated in IRIS, one possible cause is that it may actually prove to be a MEOC.

Indeed, with the National Research Council's ongoing review of the IRIS assessment development process (NRC, 2013), the focus has been devoted to ensuring that health-effect literature searches are comprehensive for the chemicals that are under examination. But the solicitation and selection processes used for IRIS to accommodate new chemicals (US EPA, 2012a) may serve to preempt consideration of chemicals for which there are no data (which necessarily might include MEOCs); the selection assessment criteria all rely on whether data already exist for a candidate chemical. A half-dozen or so criteria are used to justify selection of the comparatively few chemicals targeted for IRIS. This selection process inadvertently serves to reinforce the deflection of attention from the vast numbers of chemicals that are filtered from consideration to include in IRIS. A major outcome of this self-selection bias is that chemicals that have long been ignored (e.g., for inclusion in IRIS) risk being perpetually ignored; whether justifiably or not is unknown.

Quality control safeguards do not exist to force examination of why data for certain chemicals known to experience significant commercial usage might be totally absent from the literature. Such feedback loops could ensure that chemicals are examined for whether they lack environmental occurrence data – and therefore pose an unknown potential for exposure – and that those posing the most risk are comprehensively prioritized for inclusion. When data are indeed totally absent, a feedback loop could force conjectures or hypotheses as to whether the potential for risk might need to be examined. With this in mind, once possible MEOCs are identified, two additional lists could be created to augment databases such as IRIS. One list could comprise MEOCs that probably pose insignificant risks and another for those MEOCs meriting further examination. With access to these two additional lists, end users of databases such as IRIS would be better informed as to what the absence of a chemical from the database might actually mean. Worth noting is that some programs, such as the EPA's Unregulated Contaminant Monitoring Rule (US EPA, 2012b), do attempt to secure occurrence data for certain chemicals (albeit limited numbers) that may pose exposure risks. By continually compiling and refining lists of potential MEOCs, protocols could be established to rule them in or out as environmental contaminants and also avoid the formulation of misleading conclusions by end-users.

Failure to understand the complete scope of those chemicals that actually contaminate the environment (a subset of the universe of potential chemical contaminants) could lead to flawed life cycle sustainability assessments. This is especially true when the chemical itself is an ultimate commercial product as opposed to a by-product from other anthropogenic or natural processes. For these chemicals, sustainability proxy measures (such as carbon footprints) may not suffice in performing life cycle assessments (Laurent et al., 2012). Using APIs as the example, incomplete knowledge of the environmental occurrence of all APIs within a given therapeutic category would prevent informing rational prescribing decisions for directing the selection of those APIs that are least problematic regarding environmental impact (Daughton and Ruhoy, 2013).

In general, published data of any kind – whether positive or negative – tends to strongly influence subsequent research, which, in turn, leads to renewed focus on incrementally extending prior work. Chemicals that receive no attention (those lacking any environmental data) are at risk of eventually being relegated to obscurity. Focusing on incremental additions to existing knowledge can result in non-optimal investment of resources and lead to biased or erroneous conclusions regarding measures of risk or sustainability. Most importantly, the magnitude of a possible vulnerability in risk assessment posed by MEOCs may demonstrate that efforts focused on sustainable use of chemicals may not recognize that they are aiming at partially obscured targets – or worse yet – the wrong targets.

Positive environmental occurrence data and data of absence (levels below detection — negative data), by themselves alone, may provide a highly biased view of the potential for exposure. Positive occurrence data may perpetuate continued biased attention because of availability bias. In contrast, identifying that which is absent from the published literature serves to better define uncertainty, minimize data disparity, and improve credibility. With a framework to guard against the Matthew Effect and to reapportion focus on MEOCs, a number of positive outcomes are made possible. These include: (i) reducing unnecessary duplication of future effort by avoiding replicate studies on the same, favored monitoring targets, (ii) identifying new targets for monitoring (expanding the chemical space known to be occupied by potential stressors), (iii) assessing why MEOCs have not been previously targeted, (iv) using the incidence of MEOCs as a measure of a research field's maturity, (v) preventing misinterpretation of the absence of certain chemicals from various databases (such as IRIS) as being equivalent to a lack of any imperative for concern regarding risk, (vi) eventually using occurrence data for chemicals that used to be MEOCs to ground truth predictive models, and, most importantly, (vii) determining if specific exposure risks are being overlooked; progress with the last point can be augmented with parallel studies involving biological effects-directed studies (e.g., Blasco and Picó, 2009).

2. Methods

Establishing the absence of data for a particular API requires comprehensive examination of the published literature using search strategies that are more exhaustive, onerous, and time consuming than those that would ordinarily be used to locate positive (and sometimes negative) occurrence data. The data source used for the searches in this study is a comprehensive full-text bibliographic database maintained at the US EPA (Daughton and Scuderi, 2013) using EndNote X6 (Thomson Reuters). This database focuses exclusively on the many aspects of pharmaceuticals as environmental contaminants (one aspect of which is environmental occurrence data). Its scope and coverage are described here: http://www.epa.gov/ppcp/pdf/Synopsis-of-PPCPs.pdf. The database is updated and curated on a daily basis. Its scope spans the published archival literature, in-press articles, and the gray literature, with coverage dating back to the 1970s, which coincides with the advent of concerted study of pharmaceuticals in the environment. All documents added to the database are examined to ensure their contents are digitized; when the main bodies of documents comprised scanned images, they were digitized using Adobe Acrobat X Professional.

For each of more than 200 of the most widely prescribed or sold APIs (see the footnotes in Supplementary Table S1 for citations to the two lists used in this study — drugs still on-patent and generics), comprehensive and exhaustive queries were performed to ascertain whether environmental occurrence data existed (both positive and negative) and for which particular environmental compartments (e.g., sewage, sediments, ambient waters, biosolids, and aquatic tissues). Note that some APIs are accounted for in redundant entries among the two lists — serving as active ingredients in multiple drugs; many of the APIs in the brand-name list were also on the generic list.

In compiling these data, no attempt was made to assess data quality or the reliability of the studies; indeed, an unknown portion of these data may have been flawed. In particular, no attempt was made to assess or record the analytical limits of detection (LODs) associated with negative data (when available); the power of negative data in establishing the absence of a targeted API increases as the LODs are lowered.

Once it was clear that occurrence data (either positive or negative) for a given API could be easily located, continued searching was terminated, and the API was excluded from further consideration as a potential MEOC. Searching was continued for each API for which occurrence data could not be found. Boolean searches (with EndNote X6) used as many terms or identifiers for these APIs as possible. For those APIs lacking any data in the current PPCPs database, searches were repeated across documents newly added to the database up until the time this manuscript was submitted for publication; newly added documents are usually those that are in-press or newly published on-line.

Data published subsequent to this assessment would be expected to continually reduce the numbers of APIs found to be lacking data, which totaled 73 as of 23 May 2013 (Table S1 and its companion summary, Table 1) for the specific and limited lists used here. This reduction would probably only proceed gradually in the absence of a concerted effort to capitalize on using absence-of-occurrence data to select APIs for targeting in future monitoring studies. Continued failure to locate API occurrence data for a particular API – with increasing effort expended in searching – pointed to an ever-higher probability of absence of data. While an absence of data can never be claimed with certainty, searches were designed with the most inclusive possible criteria — for example, using unique identifiers such as the CAS registry numbers (CASRN) and the more common synonyms for API generic or trade names; in some instances, unique non-English variant spellings or unique stems were also used. The full-text document revealed by each query was then visually examined for data pertinent to environmental occurrence. Searching for terms such as “absence of data” or “lack of data” in association with a particular API can locate rare instances where the authors claimed that data had yet to be reported; this can sometimes work as a shortcut to focus searches for identifying likely MEOCs.

There are some pitfalls in digitized text searches. Full-text searching is compromised when data are masked in tables or illustrations that are imbedded as graphic images that cannot be digitized; errors also occur when optical character recognition is not faithful. This shows the importance of visually scanning the entirety of a document in which the targeted API search term is found within the narrative. Supplemental, on-line data can pose additional challenges. All search strategies will fail when the search term is misspelled in the document or when the document uses foreign spellings unknown at the time of the search.

Bona fide data of absence reflects the ability of chemical analysis to detect ever-lower levels of a chemical in a targeted environmental matrix. It means either that a targeted chemical occurs at levels below those that could be reliably detected or that residues of the chemical simply do not occur in the matrix. In contrast, the absence of data reflects a complete lack of knowledge, for whatever reason — whether the potential presence of a chemical in the environment has been

Table 1
Summary listing of potential API MEOCs.[a]

No published data (absence of data) (33 APIs)	No published data (cont'd)	Minimal published data[b] (20 APIs)	Limited published data[c] (20 APIs)
Alendronate (121268-17-5)	Olmesartan medoxomil (144689-63-4)	Chlorthalidone[d] (77-36-1)	Allopurinol (315-30-0)
Aripiprazole (129722-12-9)	Ondansetron[d] (99614-02-5)	Clonazepam[d] (1622-61-3)	Amiodarone[d] (1951-25-3)
Baclofen (1134-47-0)	Oxybutynin (5633-20-5)	Colchicine (64-86-8)	Benztropine[d] (86-13-5)
Benazepril (86541-75-5)	Phenazopyridine[d] (94-78-0)	Dicyclomine[d] (77-19-0)	Budesonide (51333-22-3)
Benzonatate (104-31-4)	Phentermine (122-09-8)	Doxazosin (74191-85-8)	Carvedilol (72956-09-3)
Buspirone[d] (36505-84-7)	Pramipexole (104632-26-0)	Etodolac[d] (41340-25-4)	Clonidine[d] (4205-90-7)
Carbidopa (28860-95-9)	Pregabalin (148553-50-8)	Formoterol (73573-87-2)	Divalproex (76584-70-8)
Cefdinir (91832-40-5)	Quinapril (85441-61-8)	Glipizide (29094-61-9)	Drospirenone (67392-87-4)
Clobetasol (25122-41-2)	Ropinirole[d] (91374-21-9)	Isosorbide (652-67-5)	Ezetimibe[d] (163222-33-1)
Cyclobenzaprine[d] (303-53-7)	Sumatriptan[d] (103628-46-2)	Memantine (19982-08-2)	Felodipine (72509-76-3)
Diphenoxylate (915-30-0)	Terazosin[d] (63590-64-7)	Mometasone furoate[d] (83919-23-7)	Finasteride[d] (98319-26-7)
Guaifenesin (93-14-1)	Tiotropium bromide (186691-13-4)	Nebivolol (99200-09-6)	Fluocinonide (356-12-7)
Hydralazine (86-54-4)	Tizanidine[d] (51322-75-9)	Nitrofurantoin (67-20-9)	Fluticasone[d] (90566-53-3)
Hydroxychloroquine[d] (118-42-3)	Topiramate (97240-79-4)	Pioglitazone[d] (111025-46-8)	Levetiracetam (102767-28-2)
Lisdexamfetamine (608137-32-2)		Quetiapine[d] (111974-69-7)	Meclizine (569-65-3)
Methocarbamol (532-03-6)		Sitagliptin (486460-32-6)	Meloxicam[d] (71125-38-7)
Montelukast[d] (158966-92-8)		Spironolactone (52-01-7)	Metoclopramide[d] (364-62-5)
Nabumetone (42924-53-8)		Tamsulosin (106133-20-4)	Norgestimate (35189-28-7)
Nitroglycerin (55-63-0)		Trazodone[d] (19794-93-5)	Nystatin (1400-61-9)
		Valacyclovir (124832-26-4)	Ramipril (87333-19-5)

[a] Data summarized from the detailed data provided in Supplemental Table S1 ("Frequently prescribed APIs lacking environmental monitoring data: (i) absence of data, (ii) data of absence, or (iii) reports restricted to trace levels") for those APIs having very limited published data on environmental occurrence. Chemical Abstracts Service Registry Numbers (CASRN) listed in parentheses after API generic names.

[b] Minimal published data = maximum of 1–2 published reports showing limited positive data.

[c] Limited published data = maximum of 3–6 published reports showing limited positive data.

[d] APIs that are among the 106 highlighted by Howard and Muir (2011) as "high production volume pharmaceuticals that have not been detected in the environment but are estimated to be persistent and/or bioaccumulative."

actively ignored (e.g., because of a lack of a suitable analytical method to support any monitoring or because of very low modeled Predicted Environmental Concentrations — PECs), the chemical has simply been overlooked or omitted from targeting (e.g., MEOCs), or any number of other, more specific reasons, including failure to report data of absence (see Table 2).

3. Results

Those APIs for which extensive occurrence data (both negative and positive) had been published were omitted from further examination at the outset of the literature searches. Out of more than 200 of the most frequently prescribed generic and trade name drugs (during 2010–2011), many have considerable published data on their environmental occurrence in a wide spectrum of matrices; some random examples include the following dozen: azithromycin, carbamazepine, ciprofloxacin, diazepam, diltiazem, fluoxetine, hydrocortisone, ibuprofen, metformin, omeprazole, paroxetine, and propranolol.

A subset of 82 APIs remained for more extensive literature evaluation (detailed in Table S1). The complete published data that support the very limited data in Table S1 derive from only 156 publications. All but three of these references were published in the last 10 years; over 71% were published since 2010 (19 in 2013, 32 in 2012, 33 in 2011, and 27 in 2010). For this subset of 82 APIs, it was not evident at the start whether any might have adequate published data. A total of nine of these 82 APIs were subsequently assessed as trending toward having occurrence data sufficient to exclude them from further literature examination; these were: clopidogrel, digoxin, duloxetine, famotidine, hydroxyzine, lansoprazole, methylprednisolone, olanzapine, and triamcinolone. The remaining 73 APIs were classified into three groups (Table 1) characterized as having: (1) no published data (absence of data), (2) minimal published data (maximum of one or two published reports showing limited positive data), and (3) limited published data (maximum of three to six published reports showing limited positive data). The respective numbers of the original group of over 200 APIs distributed among these three groups were 33, 20, and 20. Combined, these APIs composed a significant portion of the original group that comprised the most highly prescribed drugs. By extension, it seems probable that those APIs with a lack of occurrence data would tend to comprise an ever-larger portion of the API universe as less-frequently prescribed drugs are assessed. Of these APIs, 14 are noteworthy in that they had already been highlighted over as decade ago as representative of those for which concerted environmental monitoring had not yet been performed (Daughton, 2001; see Table II, therein), and they have remained untargeted in monitoring: alendronate, amiodarone, benazepril, chlorthalidone, clonazepam, cyclobenzaprine, doxazosin, glipizide, guaifenesin, pramipexole, quinapril, ropinirole, spironolactone, and terazosin.

4. Discussion

Of the 73 APIs captured in Table 1, 36% (26 APIs) are among the 106 APIs highlighted by Howard and Muir (2011) as "high production volume pharmaceuticals that have not been detected in the environment but are estimated to be persistent and/or bioaccumulative." Of these 26 APIs, 10 were among the 33 with no published occurrence data, 8 were among the 20 with minimal published data, and 8 were among the 20 with limited published data.

The lack of environmental monitoring data for the vast majority of APIs – and the overwrought focus on a select few – had first been noted over 10 years ago (e.g., Daughton, 2001, 2004b). The potential occurrence of APIs that have never been targeted for monitoring in finished drinking water has also been identified as a major unaddressed question (Daughton, 2010b).

In addition to the first direct report of APIs lacking environmental occurrence data – published in Daughton and Ruhoy (2013) and elaborated upon in Daughton (in press) – evidence of the Matthew Effect is indirectly alluded to in a study that examined published data on the presence of 203 APIs across 41 countries (Hughes et al., 2013). This latter study revealed that most of the monitoring efforts had been devoted to just 14 of the 203 APIs — a mere 7% of the total targeted.

Although the failure to consider MEOCs for targeted monitoring may often be an unconscious decision, there are instances where the selection of APIs to target for monitoring is explicitly based solely on those APIs that have been identified in prior studies (e.g., see: Tabe et al., 2010).

In the final analysis, the complete absence of occurrence data for an API may represent one extreme in a broader distribution of occurrence frequencies. For many APIs, occurrence data may populate a long tail comprising isolated reports (e.g., spread across diverse and unique environmental sample matrices) that gradually build into a

Table 2
Factors affecting the acquisition, reporting, or corroboration of negative data or absence of data for an API in the environment.[a]

Cause for negative data or absence of data	Ramification	Example or miscellaneous notes
Analytical methodology limitations:		
Analytical limits of detection (LOD) or quantitation (LOQ) not sufficiently low for API to be detected or quantified in environmental matrices.	Potent medications, even when highly prescribed, contribute relatively low, overall quantities of APIs to the environment because of their exceedingly low dosages.	Highly potent APIs (HPAPIs) (Bormett, 2008 [1 Sept]). Potent APIs include hormones, certain synthetic opioids, and glucocorticoids; fluticasone is one example.
API is problematic for routine analytical methodologies.	Analytical limitations: poor extraction efficiency; prone to matrix interferences; poor chromatography; failure to account for reversible conjugates.	API may be present but cannot be demonstrated.
API reference standard not available.	Identity of tentatively identified API cannot be verified.	API may be present but cannot be verified.
Non-homogeneity of stream flow affects sampling design.	Variable dilution of sewage and surface water flows or impact of seasonal events can impact measured levels.	Adverse weather can not only lead to diluted streams, it can also impair the effectiveness of sewage treatment or natural removal processes.
Negative data is often presented in publications with unknown or unclear analytical figures of merit.	LOD/LOQ often not reported.	High LOD can lead to erroneous negative data for APIs having substantial environmental presence.
Idiosyncrasies of drug prescribing or usage:		
Nationwide consumption data do not reflect local consumption–can confound data obtained from monitoring in locale where drug may not reflect common usage.	Local prescribing practices or customs may exclude certain APIs popular in other locales.	Most research studies involving monitoring restrict themselves to a particular geographic locale; few studies perform statistical sampling across larger regions or especially an entire nation.
Inter-country differences in drug prescribing and usage practices.	Published monitoring data for an API may come from countries where the frequency of prescribing for the parent drug is not representative of other countries.	Published monitoring data may also come from a country where the drug is readily available OTC but only available via prescription elsewhere.
Regional disease patterns may make consumption of certain medications episodic, sporadic, or irregular.	Some diseases are endemic to specific locales. Others depend on the time of year, as affected by the weather.	Epidemics or seasonal viral infections, for example, can increase the consumption of a wide variety of medications.
Poor patient compliance/adherence for certain frequently prescribed drugs can confound the interpretation of prescribing or usage data.	Poor compliance leads to leftover medications that are never consumed.	This could lead to measured environmental levels much lower than those predicted by models based on sales or prescribing data.
API manufactured in extremely low annual quantities.	Extremely low environmental loadings even under the most favorable scenarios, such as 100% excretion unchanged API.	Norelgestromin manufactured at 0.34 kg/year as of 2004 (Besse and Garric, 2009).
API newly introduced to market.	API has not experienced sufficient usage to display detectable environmental loadings.	Any of the NMEs[b] approved by FDA each year.
Idiosyncrasies of drug metabolism or excretion:		
Diurnal variations in excretion impact the statistical design of sewage sampling/monitoring.	Cyclic levels of API in discharged sewage (as well as a non-homogeneous matrix) necessitate sampling protocols that accommodate continual random and diurnal fluctuations in levels.	Levels in discharge can be affected by morning urine voids or by daily timing of doses.
Pharmacokinetics does not favor excretion of sufficient quantities of parent API.	Some APIs are extensively metabolized to inactive products, or reversible metabolic conjugates are not formed; however, environmental residues could still result from disposal of unwanted leftover drugs to sewers or trash.	Certain drugs (i.e., prodrugs) are designed to be rapidly converted to other, more active forms (which themselves may also serve as APIs in other medications).
Certain individual APIs may have multiple origins–originating from structurally different drugs.	Some APIs can be created during sewage treatment via microbial metabolism of other APIs. Some APIs are the putative agent from other prodrugs.	Canrenone can be formed from either spironolactone or canrenoate; see discussion on prodrugs (Daughton, 2013 [in press]).

Cause for negative data or absence of data	Ramification	Example or miscellaneous notes
Characteristics of transport and fate:		
API does not reside in environmental compartment targeted for monitoring.	API partitions extensively to sewage sludge or sediments (e.g., oxytetracycline).	Monitoring the wrong compartment will return negative data.
API is easily removed by conventional STPs.	API undergoes facile biodegradation or partitions extensively to sludge/biosolids.	API cannot establish sufficient environmental loading. Results in true data of absence.
API is readily removed or transformed by ambient environmental processes.	Biodegradation, sorption, sequestration, hydrolysis, photolysis.	Formation of active or inactive transformation products; possible mineralization. Results in true data of absence.
Waste streams from API manufactures or formulators can confound data obtained from receiving waters.	Levels of a select few APIs from manufacture waste streams can contribute inordinately to the total environmental loadings relative to end use.	(Larsson et al., 2007)
API overlooked or escapes notice for any number of reasons during monitoring studies:		
Newly registered NME formulated in a brand-named drug newly introduced to market.	API not yet considered or recognized for targeted monitoring.	Some NMEs gain rapid market penetration and quickly ascend the rankings of prescribing frequency or usage and therefore lag in capturing attention for targeted monitoring.
Lack of suitable analytical method.	API never targeted or targeted only in limited-scope laboratory research for demonstrating applicability of new analytical methods.	API not targeted in sufficient number of monitoring studies. Resulting data not statistically representative.
API is actively ignored as a target because it is assumed benign in the environment or assumed to have little probability of entering or persisting in the environment.	API not targeted because of fallacious assumptions regarding potential environmental significance (purported low ecotoxicity) or ability to enter the environment (e.g., structurally labile or extensively metabolized and therefore not excreted).	Knowledge regarding pharmacokinetics or environmental fate and transport may be insufficient for predicting environmental occurrence. Both bathing and disposal of leftovers to sewers, for example, could confound predicted environmental concentrations.
The Matthew Effect	Overwrought focus on those APIs subject of previous monitoring. Those APIs never targeted for monitoring continue to escape targeting.	Discussed in this paper; examples first presented in Daughton (2013 [in press]).
API detected as an unknown during analysis but subsequent identification not attempted.	API ignored as unidentified unknown during analysis. Or TICs (tentatively identified compounds) never verified in follow-up studies.	Non-targeted analysis can be time-consuming and costly. Analytical reference standard not available.
Data of absence (negative data) not reported.	Occurrence data withheld from publishing can perpetuate the absence of data.	Author does not see value in reporting negative data or journal space limitations ultimately result in its censure.
Data problematic to locate, retrieve, mine, or interpret:		
Data exist only in obscure reports; language translations not easily obtained; documents not available via the Internet.	Published occurrence data (either positive or negative) may exist but are difficult to locate or retrieve. Negative data is rarely of primary interest to investigators and is therefore not always highlighted in articles.	Negative data published in poorly accessible sources (non-native languages and gray literature not available via the Internet).
Difficulties in making literature searches comprehensive.	Difficult to accommodate all alternate spellings or unique (or arcane) synonyms for an API during literature searching; search engines are often limited with respect to full-text stemming and lemmatization capabilities. Documents with mis-spellings are not located. Even "unique" identifiers (e.g., CAS Registry Numbers - CASRN) are not a fully reliable search strategy because they can be misformatted.	Some APIs have multiple CASRN (e.g., isomers, polymorphs, salts, etc); CAS Registry Numbers are not uniformly employed in all publications.
Not all documents (or their content) are digitized. Data can remain hidden to digital searching	Full-text searching is sometimes thwarted by the fact that many documents are scanned images. For many digital documents, the search term may exist only within imbedded images.	Images are sometimes imbedded in tables, illustrations, or side-bars; these may not be amenable to digitization (e.g., angled print or characters presented as artwork may not be amenable to optical character recognition software).

Data from Besse and Garric, 2009; Bormett, 2008 and Larsson et al., 2007.
[a]Most of these issues contribute to absence of data. Very few contribute to bona fide data of absence; these are highlighted with shading.
[b]NMEs: New Molecular Entities.

higher incidence of reported positive data for other APIs. Claims for the absence of data may therefore have to be expressed in nuanced gradations — one example being the three subjective categories used in Table 1.

Data of absence for environmental presence undoubtedly results for many APIs from pharmacokinetics that favors extensive metabolism (without subsequent formation of reversible conjugates — conjugates that can undergo subsequent hydrolysis to reform the parent API). Likewise, the absence of data may sometimes result from basing decisions to ignore certain APIs as possible monitoring targets as a result of pharmacokinetics that minimizes excretion of parent API or reversible conjugates. An API may have been actively excluded from consideration for monitoring because of a low Predicted Environmental Concentration (PEC) as yielded by models based on pharmacokinetics. An adverse feedback loop may develop whereby the lack of occurrence data for an API may encourage over-reliance on predictive models.

A major problem with establishing the absence of data is the unknown extent to which studies might refrain from reporting evidence of absence (e.g., negative data). At this point it is not possible to judge whether the absence of monitoring data for certain, frequently used APIs is because of a bona fide absence of occurrence data or because of systematic failures to report data of absence. A large number of factors can contribute to data of absence (negative data) as well as to an overall absence of data; some of these are probably widely recognized (e.g., see: Oosterhuis et al., 2013). These factors are compiled in Table 2 and fall under six general categories: (i) limitations in analytical methodology or instrumentation; (ii) idiosyncrasies involved with drug prescribing or usage; (iii) idiosyncrasies of API metabolism or excretion; (iv) fate and transport characteristics; (v) failures to consider or select APIs for targeted monitoring; and (vi) difficulties in locating, mining, or interpreting published data. The Matthew Effect is provided as an example in the fifth factor (why an API has not yet been considered for targeted monitoring). The big unknown is whether the absence of data (such as caused by the Matthew Effect) explains a larger portion of missing occurrence data than does the incidence of actual, bona fide data of absence (whose contributory factors are highlighted in Table 2 by shaded cells).

Pharmacokinetic data that indicate an API is extensively excreted unchanged (and therefore indicating its propensity to enter the environment) can be used in conjunction with the absence of data to select targets for future monitoring that have a high probability of being detected. But pharmacokinetic data that shows extensive metabolism cannot be used solely in conjunction with the absence of data to indicate a lower probability of an API occurring in the environment. APIs that are extensively metabolized must first be evaluated for whether they are excreted extensively as reversible metabolic conjugates (sometimes also called interconvertable conjugates — an example of "futile metabolic cycling"), which can be reconverted to the parent API (also called deconjugation, which serves to recycle the API or aglycone). The potential role of reversible drug conjugates (including sulfate, acyl, glucuronide, and amino acid derivatives) serving as hidden reservoirs for parent APIs in the environment (Daughton, 2004a) continues to receive attention (e.g., Celiz et al., 2009). But the detailed examination of pharmacokinetic data for drugs to assess the magnitude of conjugate excretion is extremely complex (e.g., see: Trontelj, 2012) and has received comparatively little attention.

A basic problem is faced in evaluating the pharmacokinetic data submitted with commercial drug registrations. These data tend to not focus on: (i) the types and quantities of reversible conjugates of the parent API itself or (ii) the extent to which an API passes unabsorbed and non-metabolized through the gut and is directly excreted with the feces. Data regarding conjugates can be particularly confusing because it is often reported in terms of all conjugates, including those of phase I metabolites. In reality, the primary published literature needs to be thoroughly examined to determine these factors for each API; this can be extraordinarily time consuming and was not attempted for this case study. Such literature is often very limited, with the pharmacokinetics of many drugs still not being sufficiently understood. Pharmacokinetic data summaries in the secondary literature that simply state that an API is "extensively metabolized" are insufficient to rule out whether the API has potential for occurrence in the environment via excretion. The portion of excreted products involving the parent API needs to be clearly stated.

Further complications result from the numerous variables that dictate or regulate drug consumption and API metabolism and excretion. Many of the complexities involved with determining whether and how APIs enter the environment have been summarized by Daughton and Ruhoy (2009; also see Fig. 1, therein) and Daughton and Ruhoy (2013; see Table 4, therein). Pharmacokinetics can involve a complex interplay of dose regimen complexity and duration, polypharmacy (e.g., drug–drug interactions that inhibit or promote metabolism), race, age, gender, body weight, obesity, nutritional status and food intake, gut microbiota, health status (especially renal, hepatic, and gut function), circadian timing of dose, dose-form modification (e.g., splitting or crushing extended- or delayed-release tablets), and genetic factors (e.g., rapid versus poor metabolizers), among others.

Excretion data from clinical trials and research are often expressed over a defined period ranging from hours to days. Mass balance closures are rarely achieved for dose versus excreted/retained mass. Although such short-term data may be sufficient to inform pharmacodynamics, they do not reveal the extent of total excretion that ultimately occurs. With respect to the ultimate environmental loading for an excreted API, the rate of excretion is not the determinant. Instead, the determinant is the overall, cumulative excretion from a given dose (regardless of the duration of excretion).

Regardless of their pharmacokinetics, some of the APIs in Table 1 would still be expected to enter sewers directly, without undergoing metabolism. Direct entry into sewers can occur when unwanted medications are disposed into drains, and bathing can release low levels of APIs that have been excreted via sweat (Daughton and Ruhoy, 2009). More significantly, however, bathing may be the major source of release for those APIs that are used extensively or exclusively in topical medications — where high-content preparations are externally applied in significant quantities (Daughton and Ruhoy, 2009; see Table 3, therein). Among the 73 APIs listed in Table 1, for example, seven are formulated extensively in topical medications: acyclovir (formed from its prodrug valacyclovir), clobetasol, fluocinonide, fluticasone propionate, mometasone furoate, nitroglycerin, and nystatin. Targeted monitoring for at least these seven APIs may be warranted regardless of PECs derived from pharmacokinetics. Indeed, examples exist of poorly excreted APIs that are used in topical medications and which are also known to occur in the environment (Daughton and Ruhoy, 2009; see Table 4, therein).

Among the 73 APIs with potential to be classified as MEOCs (Table 1), the most prominent category comprised 33 APIs lacking any published occurrence data. These APIs composed the class with the most likelihood of being MEOCs. For these, an additional evaluation was performed to demonstrate how pharmacokinetics might be used to further evaluate these potential MEOCs for their potential to enter the environment. Those APIs that are extensively excreted unchanged or extensively excreted as reversible conjugates, and which are also notable with a complete absence of occurrence data, might have the highest potential for being MEOCs. Of these 33 APIs, a total of 20 met this criterion (Table 3); this table summarizes the data compiled in Supplemental Table S2. These 20 APIs represent those with the highest probability of being MEOCs (in addition to those used extensively in topical medications). These might therefore warrant targeting for environmental monitoring. Another category in Table 3 comprises the remaining APIs (13 in total) that are poorly excreted, which greatly limits their ability to enter the environment other than by way of direct disposal to sewers or via release of topical residues during bathing (i.e., nitroglycerin and oxybutynin).

Table 3
The 20 APIs (of 33) identified as most probable MEOCs in this study, further categorized according to one measure of their potential to enter the environment — pharmacokinetics.

	Possibly increasing probability of MEOCs →		
(−)	(+)	(++)	(+++)
Benzonatate (−?)	Aripiprazole	Alendronate	Clobetasol
Buspirone	Benazepril	Baclofen	
Cyclobenzaprine (−?)	Carbidopa	Cefdinir	
Diphenoxylate	Methocarbamol (+?)	Hydroxychloroquine	
Guaifenesin (−?)	Ondansetron (+?)	Olmesartan medoxomil	
Hydralazine	Phentermine	Phenazopyridine	
Lisdexamfetamine	Ropinirole	Pramipexole	
Montelukast (−?)	Terazosin	Pregabalin	
Nabumetone		Quinapril	
Nitroglycerin		Tiotropium bromide	
Oxybutynin		Topiramate	
Sumatriptan			
Tizanidine			

This table contains the 33 APIs that lacked any published environmental occurrence data. These are identified in Table 1. Supporting references for pharmacokinetic data relevant to each API are compiled in Supplemental Table S2 ("APIs identified as most probable MEOCs in this study, categorized according to one measure of their potential to enter the environment — pharmacokinetics").
Key:
(+++) = nearly stoichiometric release to sewers during bathing (APIs that are approved for topical use only — except for portion absorbed via the skin).
(++) = substantial portion of API (or active ingredient from prodrug form) possibly excreted unchanged or released during bathing (for topical drugs).
(+) = possibly significant quantity of API (or active ingredient from prodrug form) excreted unchanged or released during bathing (for topical drugs).
(?) = excretion of parent API either unknown or extent of reversible conjugates unknown.
(−) = little excreted unchanged or as reversible conjugates.

One criterion that can be invoked to exclude APIs from consideration as potential MEOCs pertains to certain prodrugs that lack bioactivity of their own but which release active agents that also serve as the active ingredients in other drugs. For these prodrugs, monitoring for the released API already accounts for the prodrug's entry to the environment. For example, among the narrowed list of 33 APIs in Table 3, four are prodrugs (lisdexamfetamine, nabumetone, olmesartan medoxomil, and quinapril). For one of these (i.e., lisdexamfetamine), the potential for its entry to the environment is already being inadvertently accounted for by monitoring for the active ingredient released from the prodrug — in this case, dextroamphetamine.

Finally, the case study presented here used a single selection criterion (i.e., likelihood of API excretion, both unchanged and in the form of reversible conjugates) to narrow the list of all APIs lacking occurrence data for the purpose of selecting future monitoring targets. Many additional and more difficult-to-assess factors can complicate determining the potential for an API's entry to the environment. Just because an API might be among the most frequently prescribed or sold, coupled with its being extensively excreted unchanged, does not necessarily mean that it will reach a level in the environment sufficient to be detected. Most of these responsible factors are listed in Table 2. Three worth highlighting (see: Daughton, 2010a) are: (i) prescriptions and sales do not necessarily reflect ultimate drug usage (largely because of patient non-compliance — failure to fill prescriptions or to take a full regimen of dispensed medication), (ii) drugs often experience significant geographic biases in usage and abrupt temporal changes in usage (e.g., as a function of prescribing behavior and customs, consumer demand, weather, season, epidemics), and (iii) highly potent drugs are manufactured and consumed in very low quantities. With regard to the first factor, extremely complex aspects of various human activities and behaviors surround the prescribing, dispensing, and ultimate usage of drugs. These factors can dictate which drugs are preferentially prescribed (which can reflect the knowledge, customs, and fads of prescribers as well as marketing by manufacturers) and influence patient compliance, which determines what fraction of prescriptions are eventually dispensed and how much of a dispensed drug is consumed (which, in turn, determines how much leftover drug may eventually require disposal, sometimes to sewers); some drugs are known for significantly lower patient compliance rates, especially those having unpleasant side effects or which are used to treat symptoms the patient cannot self-assess.

An API's propensity to enter the environment, however, can also be considered in conjunction with its overall hazard (e.g., see: Dong et al., 2013) which can manifest itself in many ways, including: (i) predictable acute, extreme toxicity — one example being single-dose lethality in humans (Daughton, 2010a; Daughton and Ruhoy, 2013), (ii) acute toxicity in non-target species, as witnessed by renal toxicity for certain vultures from ingesting NSAID residues (e.g., Oaks and Watson, 2011), and (iii) subtle effects (Daughton and Ternes, 1999) such as alteration of feeding, attraction, and avoidance behaviors for aquatic organisms from chronic low-level exposures (e.g., Brodin et al., 2013; Di Poi et al., 2013; Fong and Molnar, 2008; Fong and Hoy, 2012; Gaworecki and Klaine, 2008; Guler and Ford, 2010; Schultz et al., 2011; Thomas and Klaper, 2012). Additional, but more complex, aspects of toxicology also need to be considered, most prominently being interactive and additive

Table 4
Potential value of API negative occurrence data (data of absence).

Data of absence: value, utility, advantage	Explanation
Guide the selection of APIs to be targeted in environmental monitoring	APIs having insufficient data of absence (i.e., absence of data) can be cross-checked with PECs (Predicted Environmental Concentrations) to indicate their possibility of occurrence (and need for further monitoring). Data of absence reported only in particular environmental compartments may direct the need for monitoring in other compartments.
Justify the exclusion of certain APIs from future monitoring	Helps avoid duplication of effort and expenditure of additional resources; narrows the universe of APIs requiring further examination (e.g., ecotoxicity); but attention must still be paid to special circumstances (one example being the waters receiving treated waste streams from a manufacturer).
Guide medical prescribing decisions for selection of medications having a low probability of environmental impact	APIs having sufficient data of absence do not enter the environment as a result of intended end-use (e.g., via excretion or bathing) or by way of disposal of leftovers to sewers.
Ground-truth predictive models; help avoid over-reliance on predictive models	PECs indicating low probability of occurrence should be corroborated with sufficient data of absence or questioned with the absence of data.
Help guide design of environmentally benign drugs	APIs having sufficient data of absence may have structural attributes that reduce excretion or promote environmental transformation/degradation; these APIs increase the power of structure activity relationships.
Avoid the Matthew Effect	Dedicating some emphasis to acquiring and reporting negative data may help avoid an over-wrought and biased focus on a limited subset of potential API environmental contaminants and reduce the incidence of MEOCs.
Improves completeness of monitoring data (weight of evidence), providing a more representative picture of overall API occurrence in the environment	Censoring data of absence serves to bias the overall picture of API occurrence in the environment; bias toward positive findings of API occurrence in monitoring studies (suppression of negative findings) reduces the overall value of monitoring.

toxicity resulting from exposure to multiple APIs, perhaps all at individual levels that may be substantially below their known no-effect levels.

As an addendum reflecting the progress in the reporting of new data, since the cut-off date of the examination reported here (23 May 2013), an unusually large monitoring study (Chen et al., in press) reports on the tentative occurrence of 11 of the APIs compiled in Table 1. While this new study did not provide confirmation of the structural identities, each of these 11 APIs were reported in the sediments of at least one of three rivers in China at levels sufficient to give signal-to-noise ratios greater than 3 (amiodarone, *hydroxychloroquine*, and *ondansetron*) or 10 (*benazepril*, *buspirone*, *quinapril*, *tizanidine*, quetiapine, trazodone, glipizide, and clonidine); the six APIs shown in italics had not yet been reported in the literature, and the remaining five had only minimal or limited studies supporting their environmental occurrence. Although the gradual addition of new data such as these will serve to eventually erode the existing numbers of MEOCs, new candidates are continually added with newly approved drugs and with revised compilations of most-frequently prescribed drugs.

5. Future research and considerations

Several decades of monitoring studies have established the presence of hundreds of APIs in a wide variety of environmental compartments and matrices, including: treated and raw sewage (including sludge and biosolids), animal manure, landfills, surface waters, ground waters, drinking waters, marine environments, sediments, tissues of crops and native vegetation, tissues of aquatic and occasionally terrestrial wildlife, manufacturer wastewaters, and air (e.g., see: Fatta-Kassinos et al., 2011; Heeb et al., 2012; Howard and Muir, 2011; Lapworth et al., 2012; Petrovic et al., 2008; Thomas and Langford, 2010); these hundreds, however, represent but a small percentage of those APIs in active use (see: Huang et al., 2011). These data are published in a bewildering spectrum of disparate journals, dissertations, books, reports, and other gray literature. Despite these published occurrence data, however, no attempt has been made to compile them in a publically accessible database. With the first comprehensive database of extant APIs, now assembled in the National Institutes of Health (NIH) Chemical Genomics Center's Pharmaceutical Collection (NPC) (Huang et al., 2011), an opportunity exists (Daughton, in press) to crosswalk this comprehensive list of known APIs with those that have been reported in the published environmental monitoring projects. This could provide a real-time perspective on which APIs have evaded attention — the MEOCs. While construction of a comprehensive database on API occurrence in the environment from the published literature would be a major undertaking, it would be amenable to compilation and curation by crowdsourcing.

Likewise, it is also surprising that the yearly lists of the most widely used drugs (e.g., APIs most frequently prescribed, such as those used for Table S1) are not periodically evaluated for those APIs that have not yet been detected in the environment. These publically available lists can change dramatically on a yearly basis. In real time, prescribing and usage are in a constant state of flux as a result of many influences (e.g., see: Table 2). Access to real-time, geographic prescribing or sales data, which is largely proprietary in the U.S. and available only via subscription, would greatly help in maintaining dynamic lists of APIs to examine for MEOC candidates. Examples of geographic discrepancies, often resulting from prescribing customs and the incidence of disease, can be seen with the data published by Express Scripts (Cox et al., 2008) and in the work reviewed by Wangia and Shireman (in press). Examples of geographic usage data derived from direct sewage monitoring are most extensive for illicit drugs; these studies reveal large geographic discrepancies in relative usage patterns (e.g., Nefau et al., 2013; Thomas et al., 2012). Also note that some medications are not approved for use in all countries, and certain drugs can be withdrawn from markets in some countries but not others.

While the frequency of prescribing does not directly translate to the mass of the API prescribed or consumed, it does translate into those APIs whose end use probably results in better market penetration and therefore higher probabilities for detection in wastewaters and the environment; these lists only represent but a small percentage (perhaps less than 5%) of the total numbers of APIs in use. APIs lacking occurrence data should be further investigated to determine the cause of the absence of data — whether unreported APIs have been overlooked (casualties of the Matthew Effect) or actively ignored (for example, because of difficulty in analysis). Evaluation of published PEC and PNEC (predicted no-effect concentration) data for each API can serve as an initial filter to select possible MEOCs to further investigate.

With respect to pollution prevention, it is worth noting that not only can the selection of potential API monitoring targets be guided by the absence of data (e.g., MEOCs), but also drug prescribing decisions could be guided by data of absence. Those APIs whose published environmental occurrence data consistently show unmeasurable levels in a sufficiently wide range of matrices could possibly be classified as environmentally benign. Within a particular therapeutic class, and with therapeutic efficacies being similar, those APIs with negligible environmental footprints could be favored for prescribing. For this reason, it would be useful to verify which of the highly prescribed APIs lacking any type of environmental monitoring data actually do have negligible environmental presence (consistent data of absence). Other advantages are associated with data of absence (Table 4), the incidence of which can be maximized by identifying MEOCs and ruling them in or out as not occurring in the environment.

Finally, the case study presented here for APIs could be extended to other classes of chemicals to determine whether MEOCs reveal similar biases in establishing the scope of environmental monitoring programs. With an active focus on MEOCs, a different perspective could be gained for the prioritization of environmental contaminants when designing targeted monitoring programs. In particular, chemicals that are potential MEOCs should be prominently highlighted to ensure that they do not permanently escape notice. Perhaps more important is advancement of our understanding of how MEOCs come about and what protocols could be established to avoid their proliferation.

Conflict of interest

The author has no conflict of interest.

Acknowledgments

U.S. EPA Notice: The views expressed in this article are those of the author and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. This work was supported by ORD's Pathfinder Innovation Program, which was launched in October, 2010. The assistance of MST Scuderi (SEEP, U.S. EPA, Las Vegas) in maintenance of the EndNote bibliographic database is greatly appreciated.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.scitotenv.2013.06.111.

References

Besse J-P, Garric J. Progestagens for human use, exposure and hazard assessment for the aquatic environment. Environ Pollut 2009;157:3485–94.
Blasco C, Picó Y. Prospects for combining chemical and biological methods for integrated environmental assessment. TrAC Trends Anal Chem 2009;28:745–57.

Bormett D. High-potency APIs: containment and handling issues. Pharm Technol 2008. [1 Sept] [Available on: http://www.pharmtech.com/pharmtech/Ingredients/High-Potency-APIs-Containment-and-Handling-Issues/ArticleStandard/Article/detail/548824. (accessed 21 May 2013)].

Brodin T, Fick J, Jonsson M, Klaminder J. Dilute concentrations of a psychiatric drug alter behavior of fish from natural populations. Science 2013;339:814–5.

Buffler PA. The evaluation of negative epidemiologic studies: the importance of all available evidence in risk characterization. Regul Toxicol Pharmacol 1989;9:34–43.

Celiz MD, Tso J, Aga DS. Pharmaceutical metabolites in the environment: analytical challenges and ecological risks. Environ Toxicol Chem 2009;28:2473–84.

Chen YS, Yu S, Hong YW, Lin QY, Li HB. Pharmaceutical residues in tidal surface sediments of three rivers in southeastern China at detectable and measurable levels. Environ Sci Pollut Res 2013. http://dx.doi.org/10.1007/s11356-013-1871-y. [in press].

Cox ER, Mager D, Weisbart E. Geographic variation trends in prescription use: 2000 to 2006. St. Louis, MO: Express Scripts, Inc.; 200823 [Available on: http://www.express-scripts.com/research/research/archive/docs/geoVariationTrends.pdf. (accessed 21 May 2013)].

Daughton CG. Pharmaceuticals and personal care products in the environment: overarching issues and overview. In: Daughton CG, Jones-Lepp TL, editors. Pharmaceuticals and personal care products in the environment: scientific and regulatory issues. Washington, DC: American Chemical Society; 2001. p. 2–38. [chapter 1].

Daughton CG. Non-regulated water contaminants: emerging research. Environ Impact Assess Rev 2004a;24:711–32.

Daughton CG. PPCPs in the environment: future research — beginning with the end always in mind. In: Kümmerer K, editor. Pharmaceuticals in the environment — sources, fate, effects and risks. Berlin, Germany: Springer; 2004b. p. 463–96. [chapter 33].

Daughton CG. Drugs and the environment: stewardship & sustainability. Las Vegas, NV: National Exposure Research Laboratory, Environmental Sciences Division, US EPA; 2010a [NERL-LV-ESD 10/081, EPA/600/R-10/106, 196 pp. Available on: http://www.epa.gov/nerlesd1/bios/daughton/APM200-2010.pdf. (accessed 21 May 2013)].

Daughton CG. Pharmaceutical ingredients in drinking water: overview of occurrence and significance of human exposure. In: Halden RU, editor. Contaminants of emerging concern in the environment: ecological and human health considerations. Washington, DC: American Chemical Society; 2010b. p. 9–68. [chapter 2. Available on: http://www.epa.gov/esd/bios/daughton/ACS-SS1048-2010.pdf. (accessed 21 May 13)].

Daughton CG. Pharmaceuticals in the environment: sources and their management. In: Barcelo D, Perez S, Petrovic M, editors. Analysis, fate and removal of pharmaceuticals in the water cycle, Wilson & Wilson's comprehensive analytical chemistry. Elsevier; 2013. [in press].

Daughton CG, Ruhoy IS. Environmental footprint of pharmaceuticals: the significance of factors beyond direct excretion to sewers. Environ Toxicol Chem 2009;28:2495–521.

Daughton CG, Ruhoy IS. Lower-dose prescribing: minimizing "side effects" of pharmaceuticals on society and the environment. Sci Total Environ 2013;443:324–37.

Daughton CG, Scuderi MST. Pharmaceuticals and Personal Care Products (PPCPs): relevant literature. Las Vegas, NV: U.S. Environmental Protection Agency; 2013 [(a comprehensive bibliographic database of literature references; first implemented 19 February 2008). Available on: http://www.epa.gov/ppcp/lit.html. (accessed 21 May 2013)].

Daughton CG, Ternes TA. Pharmaceuticals and personal care products in the environment: agents of subtle change? Environ Health Perspect 1999;107(Suppl. 6):907–38.

Di Poi C, Darmaillacq A-S, Dickel L, Boulouard M, Bellanger C. Effects of perinatal exposure to waterborne fluoxetine on memory processing in the cuttlefish *Sepia officinalis*. Aquat Toxicol 2013;132–133:84–91.

Dong Z, Senn DB, Moran RE, Shine JP. Prioritizing environmental risk of prescription pharmaceuticals. Regul Toxicol Pharmcol 2013;65:60–7.

Escher BI, van Daele C, Dutt M, Tang JY-M, Altenburger R. Most oxidative stress response in water samples comes from unknown chemicals: the need for effect-based water quality trigger values. Environ Sci Technol 2013;47:7002–11. http://dx.doi.org/10.1021/es304793h.

Fatta-Kassinos D, Meric S, Nikolaou A. Pharmaceutical residues in environmental waters and wastewater: current state of knowledge and future research. Anal Bioanal Chem 2011;399:251–75.

Fong PP, Hoy CM. Antidepressants (venlafaxine and citalopram) cause foot detachment from the substrate in freshwater snails at environmentally relevant concentrations. Mar Freshw Behav Physiol 2012;45:145–53.

Fong P, Molnar N. Norfluoxetine induces spawning and parturition in estuarine and freshwater bivalves. Bull Environ Contam Toxicol 2008;81:535–8.

Gaworecki KM, Klaine SJ. Behavioral and biochemical responses of hybrid striped bass during and after fluoxetine exposure. Aquat Toxicol 2008;88:207–13.

Grandjean P, Eriksen M, Ellegaard O, Wallin J. The Matthew effect in environmental science publication: a bibliometric analysis of chemical substances in journal articles. Environ Health 2011;10:96.

Gray GM, Cohen JT. Policy: rethink chemical risk assessments. Nature 2012;489:27–8.

Guler Y, Ford AT. Anti-depressants make amphipods see the light. Aquat Toxicol 2010;90:397–404.

Heeb F, Singer H, Pernet-Coudrier B, Qi W, Liu H, Longrée P, et al. Organic micropollutants in rivers downstream of the megacity Beijing: sources and mass fluxes in a large-scale wastewater irrigation system. Environ Sci Technol 2012;46:8680–8.

Holzer A. Do pharmaceuticals in the environment present an investment risk? In: Kümmerer K, Hempel M, editors. Green and Sustainable Pharmacy. Berlin Heidelberg, Germany: Springer-Verlag; 2010. p. 287–92. [chapter 19].

Howard PH, Muir DCG. Identifying new persistent and bioaccumulative organics among chemicals in commerce II: pharmaceuticals. Environ Sci Technol 2011;45:6938–46.

Huang R, Southall N, Wang Y, Yasgar A, Shinn P, Jadhav A, et al. The NCGC pharmaceutical collection: a comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. Sci Transl Med 2011;3:1–16.

Hughes SR, Kay P, Brown LE. Global synthesis and critical evaluation of pharmaceutical data sets collected from river systems. Environ Sci Technol 2013;47:661–77.

IOM. Environmental decisions in the face of uncertainty. Committee on decision making under uncertainty, board on population health and public health practice, Institute of Medicine [IOM]. The National Academies Press; 2013. p. 260. [Available on: http://www.nap.edu/catalog.php?record_id=12568. (accessed 21 May 2013)].

Lapworth DJ, Baran N, Stuart ME, Ward RS. Emerging organic contaminants in groundwater: a review of sources, fate and occurrence. Environ Pollut 2012;163:287–303.

Larsson DGJ, de Pedro C, Paxeus N. Effluent from drug manufactures contains extremely high levels of pharmaceuticals. J Hazard Mater 2007;148:751–5.

Laurent A, Olsen SI, Hauschild MZ. Limitations of carbon footprint as indicator of environmental sustainability. Environ Sci Technol 2012;46:4100–8.

Merton R. The Matthew effect in science: the reward and communication systems of science are considered. Science 1968;159:56–63.

Nefau T, Karolak S, Castillo L, Boireau V, Levi Y. Presence of illicit drugs and metabolites in influents and effluents of 25 sewage water treatment plants and map of drug consumption in France. Sci Total Environ 2013;461–462:712–22.

NRC. National Research Council Review of the IRIS assessment development process. US Environmental Protection Agency; 2013 [Available on: http://www.epa.gov/iris/iris-nrc.htm. (accessed 21 May 2013)].

Oaks JL, Watson RT. South Asian vultures in crisis: environmental contamination with a pharmaceutical. In: Elliott JE, Bishop CA, Morrissey CA, editors. Wildlife ecotoxicology: forensic approaches. Emerging topics in ecotoxicologyNew York, NY: Springer; 2011. p. 413–41. [chapter 14].

Oosterhuis M, Sacher F, ter Laak TL. Prediction of concentration levels of metformin and other high consumption pharmaceuticals in wastewater and regional surface water based on sales data. Sci Total Environ 2013;442:380–8.

Petrovic M, Radjenovic J, Postigo C, Kuster M, Farre M, de Alda M, et al. Emerging contaminants in waste waters: sources and occurrence. In: Barceló D, Petrović M, editors. The handbook of environmental chemistry. Emerging contaminants from industrial and municipal waste: occurrence, analysis and effects. Berlin Heidelberg, Germany: Springer-Verlag; 2008. p. 1–35.

Sass J, Rosenberg D. The delay game: how the chemical industry ducks regulation of the most toxic substances. October, New York, NY: Natural Resources Defense Council (NRDC); 201123 [Available on: http://www.nrdc.org/health/files/IrisDelayReport.pdf. (accessed 21 May 2013)].

SCENIHR, SCCS, SCHER. Preliminary report on addressing the new challenges for risk assessment. 8 October, Scientific Committee on Consumer Safety (SCCS), Scientific Committee on Health and Environmental Risks (SCHER), and Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR). European Commission; 2012. p. 154. [Available on: http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_037.pdf. (accessed 21 May 2013)].

Schultz MM, Painter MM, Bartell SE, Logue A, Furlong ET, Werner SL, et al. Selective uptake and biological consequences of environmentally relevant antidepressant pharmaceutical exposures on male fathead minnows. Aquat Toxicol 2011;104:38–47.

Tabe S, Yang P, Zhao X, Hao C, Seth R, Schweitzer L, et al. Occurrence and removal of PPCPs and EDCs in the Detroit River watershed. Water Pract Technol 2010;5:1. http://dx.doi.org/10.2166/wpt.2010.015.

Tang JYM, McCarty S, Glenn E, Neale P, Warne MSJ, Escher BI. Mixture effects of organic micropollutants present in water: towards the development of effect-based water quality trigger values for baseline toxicity. Water Res 2013;47:3300–14.

Thomas MA, Klaper RD. Psychoactive pharmaceuticals induce fish gene expression profiles associated with human idiopathic autism. PLoS One 2012;7:e32917.

Thomas KV, Langford KH. Point sources of human pharmaceuticals into the aquatic environment. In: Kümmerer K, Hempel M, editors. Green and sustainable pharmacy. Berlin Heidelberg, Germany: Springer-Verlag; 2010. p. 211–23. [chapter 14].

Thomas KV, Bijlsma L, Castiglioni S, Covaci A, Emke E, Grabic R, et al. Comparing illicit drug use in 19 European cities through sewage analysis. Sci Total Environ 2012;432:432–9.

Trontelj J. Quantification of glucuronide metabolites in biological matrices by LC–MS/MS. In: Prasain JK, editor. Tandem mass spectrometry — applications and principles. Manhattan, NY: InTech; 2012. p. 531–58. [chapter 22].

US GAO. Toxic substances: EPA has increased efforts to assess and control chemicals but could strengthen its approach. GAO-13-249, United States Government Accountability Office; 201344 [Available on: http://www.gao.gov/assets/660/653276.pdf. (accessed 21 May 2013)].

US EPA. Integrated Risk Information System (IRIS): frequent questions. US Environmental Protection Agency; 2012a [Available on: http://www.epa.gov/iris/help_ques.htm#howsub. (accessed 21 May 2013)].

US EPA. Occurrence data: accessing unregulated contaminant monitoring data. US Environmental Protection Agency; 2012b [Available on: http://water.epa.gov/lawsregs/rulesregs/sdwa/ucmr/data.cfm. (accessed 21 May 2013)].

US EPA. Chemical Data Reporting. US Environmental Protection Agency; 2013a [Available on: http://www.epa.gov/oppt/cdr/index.html. (accessed 21 May 2013)].

US EPA. Integrated Risk Information System (IRIS). US Environmental Protection Agency; 2013b [Available on: http://www.epa.gov/iris/index.html. (accessed 21 May 2013)].

Wangia V, Shireman TI. A review of geographic variation and Geographic Information Systems (GIS) applications in prescription drug use research. Res Social Adm Pharm 2013. http://dx.doi.org/10.1016/j.sapharm.2012.11.006. [in press].