

ISPE RESPONSE TO FDA QUESTIONS

Standardization and Evaluation of Risk Evaluation and Mitigation Strategies (REMS)—Public Meeting on July 25-26, 2013

The International Society for Pharmacoepidemiology (ISPE) is pleased to have the opportunity to offer its perspective and suggestions, and submits for your consideration the following response to the FDA questions relating to the public workshop REMS Standardization and Evaluation Public Meeting held July 26, 2013. ISPE comments are focused on question E2 from the FDA Notice of Public Hearing and Request for Comments Standardizing and Evaluating REMS, Vol. 78, No. 99, May 22, 2013.

ISPE is an international, nonprofit, professional membership organization dedicated to promoting the health of the public by advancing pharmacoepidemiology, the science that applies epidemiological approaches to studying the use, effectiveness, values and safety of pharmaceuticals. ISPE is firmly committed to providing an unbiased scientific forum to the views of all parties with interests in drug, biologics, and devices development, delivery, use, costs and value, adverse and beneficial effects, and therapeutic risk management. Moreover, the Society provides an international forum for the open exchange of scientific information among academia, government, and industry and for the development of policy, education and advocacy for the fields of pharmacoepidemiology, including including such areas as pharmacovigilance, drug utilization research, comparative effectiveness research, and therapeutic risk management.

The Society's more than 1,500 members represent 45 countries. ISPE members work in academic institutions, the pharmaceutical industry, service providers, government agencies, and non-profit and for-profit private organizations and institutions. ISPE members are researchers with background and training in epidemiology, biostatistics, medicine, public health, nursing, pharmacology, pharmacy, law, and health economics.

Our response is based on input from members of the International Society of Pharmacoepidemiology and was prepared by the Benefit Risk Assessment, Communication and Evaluation (BRACE) Special Interest Group, which includes representatives from

Endorsed by the ISPE Board of Directors, March 5, 2014

industry, service providers, academia, and government in North America and Europe. We thank the FDA for the opportunity to provide comments on the specific questions regarding REMS standardization and evaluation.

Lead Authors, BRACE SIG

Meredith Smith, PhD, MPA

Elaine Morrato, DrPH, MPH, CPH

Stephen Sun, MD, MPH

Alicia Gilsonan, RPh, PhD

Rachael DiSantistefano, MS, PhD

Terri Madison, PhD, MPH

Paul Coplan, ScD, MBA

David Neasham, Ph.D, MSc, MFPH

TABLE OF CONTENTS

| | |
|---|-----------|
| QUESTION A | 1 |
| 1. What are the strengths and weaknesses of using/adapting existing healthcare intervention assessment framework(s) for the purpose of drafting guidance for future REMS assessments? | 1 |
| QUESTION B | 4 |
| 1. Based on your experience/expertise: Do the domains of Reach, Effectiveness, Adoption, Implementation and Maintenance, including burden/access and root causes of program failure, cover the scope of domains of interest for REMS in order to assess whether the program is meeting its goals and/or areas for program modification? If not, what other frameworks or individual domains should be considered? | 4 |
| 2. Based on your experience/expertise: Are there specific existing or emerging methodologies (e.g., DUR, claims analysis, audits, observational studies, pre-post studies, etc) that should be considered to collect any or all of these assessment domains?..... | 8 |
| 3. What metric(s) (i.e., numerator, denominator, value, units) could be used as a standard for assessing reach, effectiveness, adoption, implementation and maintenance? For each recommended metric, what data sources are available to collect such information? Any commentary around this? | 14 |
| 4. What data analysis and statistical methods are recommended for periodically and serially analyzing REMS performance over time? (How) can target performance thresholds be established and used to determine how well the program is meeting its goal(s)? Any commentary around this? | 17 |
| QUESTION C | 20 |
| 1. Does the working group have any additional recommendations for submission to the docket? | 20 |

During the FDA Public Meeting “REMS Standardization and Evaluation” held on July 26, 2013, several questions were posed by the FDA in regard to using a standard framework to guide the evaluation of REMS programs. One specific approach, the RE-AIM framework, was referenced in particular. RE-AIM was developed by Russell Glasgow and colleagues (1999). It conceptualizes the public health impact of an intervention as a function of five factors: reach, efficacy/effectiveness, adoption, implementation and maintenance. While RE-AIM addresses many salient aspects of evaluation, it is not the only comprehensive evaluation framework available in the published literature (see below for a list of other frameworks). Program evaluation is an evolving science and evaluation planners should select the framework that is most appropriate for the particular program at hand.

Specific questions were posed about the RE-AIM framework to the BRACE SIG in order to gather input from the diverse perspectives of the SIG membership. These questions and our responses are presented below. In particular, we discuss the value of using a framework for evaluation purposes, reference the range of currently available evaluation frameworks, highlight the types of domains that should be addressed in REMS evaluations, and describe possible methodologies and metrics to use in assessing those domains.

Question A

1. *What are the strengths and weaknesses of using/adapting existing healthcare intervention assessment framework(s) for the purpose of drafting guidance for future REMS assessments?*

ISPE does not advocate the application of any particular framework. Rather, we suggest that sponsors use a framework to guide their REMS evaluation, selecting from among the many frameworks that are available. Sponsors should provide a rationale for their particular selection.

An array of healthcare intervention assessment frameworks is applicable to the planning, implementation, and assessment of REMS programs. For example, the National Cancer Institute (NCI) has listed five such frameworks that have had extensive empirical application (NCI, 2012):

- Canadian Institutes for Health Research Model of Knowledge Translation
- Consolidated Framework for Implementation Research (CFIR);
- Interactive Systems Framework: A Practical, Robust Implementation and Sustainability Model (PRISM)
- Precede-Proceed Model
- Aday & Anderson framework (1999)

- Donabedian Quality of Care framework (1980)
- Reach, Effectiveness, Adoption, Implementation, Maintenance (RE-AIM)

The value of frameworks from a REMS perspective is that they are grounded in the social and behavioral sciences, focus on issues of dissemination and implementation of public health interventions under real-world circumstances, and are intended to guide the development of health care interventions involving behavioral change at both the individual and system levels.

Currently, there is no consensus within the field regarding the preferability of one framework over the other. Across these frameworks, however, there are a number of core domains that they all share, although there is not a single framework that encompasses all domains. These domains address the degree to which the program has the following attributes:

- Evidence-based
- Stakeholder-centric
- Implemented with fidelity
- Reaches a significant portion of the intended audience and is adopted by the targeted health care settings
- Effective
- Sustainable (i.e., shows evidence of being sustained under real-world conditions).

The following abilities of such frameworks contribute to their strength (Tabak et al., 2012):

- Enhance the dissemination and effectiveness of health behavior interventions (such as REMS) and increase their likelihood for success by helping to focus the interventions on the essential processes of behavioral change, which can be quite complex
- Ensure that interventions are designed for dissemination purposes
- Ensure that essential implementation strategies are included
- Enhance interpretability of the findings to a range of audiences, including regulators, industry, health care providers, media and other lay audience, and patients.
- Provide a scientifically rigorous structure and consistency to the evaluation process, thereby ensuring a stronger scientific basis for REMS program design in the future
- Enable the FDA and sponsors to compare REMS metrics and outcomes across REMS programs

- Ground findings in the larger body of existing empirical research that has been conducted in the scientific field of health care intervention dissemination and implementation.

The following are limitations of selecting only one framework to apply to all REMS situations:

- A single universal framework is unlikely to have all core domains to cover all possible REMS programs
- Flexibility in choosing an approach to assess individual REMS programs is limited
- Value of benchmarking REMS programs between different product classes and therapeutic use is limited or not relevant

In summary, the application of a dissemination and implementation framework can provide a comprehensive and systematic approach to the design and implementation of REMS strategies and assessments. A guide for selecting an appropriate framework for a specific REMS is available from the Veterans Administration QUERI Enhancing Implementation program for fostering implementation in health services (Damschroder et al., 2009).

References

Aday, LA, et al, A framework for assessing the effectiveness, efficiency, and equity of behavioral healthcare. *Am J Manag Care*, 1999. **5**: p. SP25-SP44.

Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. 2009 Aug 7;4:50.

Donabedian, A., Explorations in quality assessment and monitoring: the definition of quality and approaches to its assessment. 1980, Ann Arbor, MI: Health Administration Press.

Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Pub Health* 1999, 89:1322-1327.

National Cancer Institute. Dissemination and implementaiton models and frameworks. 2012. Available at:
<http://cancercontrol.cancer.gov/is/fundingresources/modelsandframeworks.html>.
Accessed 7 January 2014.

Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. *Am J Prev Med.* 2012 Sep;43(3):337-50.

Question B

1. *Based on your experience/expertise:*

Do the domains of Reach, Effectiveness, Adoption, Implementation and Maintenance, including burden/access and root causes of program failure, cover the scope of domains of interest for REMS in order to assess whether the program is meeting its goals and/or areas for program modification? If not, what other frameworks or individual domains should be considered?

While RE-AIM (Reach, Effectiveness, Adoption, Implementation and Maintenance) has many domains of interest, other frameworks also provide additional domains that are relevant from the perspective of a REMS program evaluation. After reviewing several possible frameworks (CDC, 2012; Chen, 1990; Damschroder, 2009; Dearing, 2013, Fischhoff, 2011; FDA, 2005; FDA, 2009; Glanz, 2008; Glasgow, 1999; Rogers, 2003; Tabak, 2012) we recommend the core set of domains in Table 1 for consideration. The proposed seven domains, which encompass the five specified domains in RE-AIM, represent the minimum set of domains that should be considered for any REMS evaluation. However, any individual REMS program evaluation may incorporate additional domains (derived from any of the previously referenced frameworks, for example) as deemed relevant or necessary.

Core Set of Domains to Consider for the Purposes of REMS Evaluation

We propose that the core set of domains listed in Table 1 be considered when designing, implementing and evaluating a REMS program. These domains are listed sequentially in the order in which they should be addressed when planning a REMS evaluation.

Table 1. Proposed Core Set of Domains to be Assessed

| Domain | Definition |
|--|---|
| Evidence-based program components* | <p>Extent to which</p> <ul style="list-style-type: none"> ▪ REMS program components incorporate interventions that have been shown to be empirically effective by utilizing components and/or materials that have been pretested and validated (e.g., undergone formative and summative testing) ▪ REMS program components have undergone piloting (and beta testing if applicable) ▪ REMS program is sufficiently “adaptable” or robust that the core program components can be delivered across different health care settings, health care systems, etc. without compromising effectiveness (this is typically based on a qualitative assessment) |
| Stakeholder-centered program components* | <p>Extent to which</p> <ul style="list-style-type: none"> ▪ REMS program components have been identified based on input from targeted stakeholder group(s) ▪ REMS program effectiveness measures were developed based on input from targeted stakeholder group(s) as to what measures were relevant from their perspective |
| Implementation** | <p>Extent to which the REMS program was delivered as intended. Specifically, the degree to which</p> <ul style="list-style-type: none"> ▪ Implementers of the program were trained to implement the program ▪ Program recipients were trained (if applicable) ▪ Program components (including program materials) were developed as planned (e.g., content was consistent with original plan) ▪ Program components (including program materials) were delivered as planned, that is, in the planned frequency and amount (or “dose”) |
| Reach and adoption** | <p>Extent to which the REMS program was delivered to the target participants and was embraced by the targeted settings (e.g., pharmacies, in- or outpatient sites):</p> <ul style="list-style-type: none"> ▪ “Reach”**: Number and percentage of eligible patients, caregivers, and health care professionals (HCPs) who enrolled/joined/participated in the program, and the descriptive characteristics of those participants (e.g., sociodemographics, health and clinical status, psychosocial characteristics, and practice characteristics [for HCPs]). ▪ “Adoption”: Number and percentage of eligible health care settings/sites/practitioners that participated in the program and/or implemented the desired behavior and descriptive characteristics of those settings or individuals (e.g., size of case load, in- or out-patient, insurance coverage mix). ▪ “Access”: Extent to which those targeted to receive the program encountered barriers to accessing the program, the timeliness with which they were able to access the program, types of barriers encountered in accessing the program, and factors that served to promote access. [Note: this may be difficult to assess for some programs.] |

| Domain | Definition |
|----------------------------------|---|
| Effectiveness** | <p>Extent to which the REMS program achieved targeted process and outcome goals. Specifically, the extent to which the targeted threshold of change was achieved in the following areas (not all will necessarily be applicable to each REMS program):</p> <ol style="list-style-type: none"> 1. Participant awareness of safe and appropriate use of drug and drug-related risks 2. Participant knowledge of safe and appropriate use of drug and drug-related risks 3. Participant attitudes regarding drug, its risks, and how to safely and appropriately use it 4. Participant behavioral intent (patients, caregivers, HCPs, etc.) regarding safe and appropriate use of drug 5. Participants' actual behavior regarding safe and appropriate use of drug 6. Patient outcomes, including disease-specific quality-of-life indicators, and clinical endpoints 7. Unintended outcomes <p>[1-5 are process measures; 6 and 7 are outcome measures.]</p> |
| Maintenance and sustainability** | <p>Extent to which the REMS program</p> <ul style="list-style-type: none"> ▪ Was delivered as planned consistently over time across participating target groups (e.g., patients, HCPs) ▪ Was "institutionalized" or integrated into the participating health care settings and sustained over time |
| Resource utilization* | <p>Amount of resources incurred in launching and maintaining the REMS program over time from the perspective of the implementing health care organization (e.g., physician office, pharmacy, individual health care provider). Specifically,</p> <ul style="list-style-type: none"> ▪ Number of staff needed to implement and run the program over time ▪ Extent to which the program fits within the existing work flow or processes within the implementing sites/settings ▪ Extent to which program participation involves additional time, money, equipment, and special training |

* New domains, not covered in RE-AIM; ** Domains covered in RE-AIM with modified definitions.

References

Centers for Disease Control and Prevention (CDC). A framework for program evaluation. 2012. Available at: <http://www.cdc.gov/eval/framework/index.htm>. Accessed January 7, 2014.

Chen HT. Theory-driven evaluations. Newbury Park, CA: Sage Publications; 1990.

- Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.* 2009 Aug 7;4:50.
- Dearing JW, Smith DK, Larson RS, Estabrooks CA. Designing for the diffusion of a biomedical intervention. *Am J Prev Med.* 2013;44(IS2):S70-S76.
- Fischhoff B, Brewer NT, Downs JS (editors). *Communicating risks and benefits: an evidence-based user's guide.* US Department of Health and Human Services, Food and Drug Administration, Risk Communication Advisory Committee and consultants. 2011. Available at: <http://www.fda.gov/AboutFDA/ReportsManualsForms/Reports/ucm268078.htm>. Accessed January 7, 2014.
- Food and Drug Administration (FDA). Guidance for industry. Development and use of risk minimization action plans. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER); 2005. Available at: www.fda.gov/downloads/RegulatoryInformation/guidances/ucm126830.pdf. Accessed January 7, 2014.
- Food and Drug Administration (FDA). Guidance for industry. Label comprehension studies for nonprescription drug products. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER); 2009. Available at: www.fda.gov/downloads/RegulatoryInformationDrugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM143834.pdf. Accessed January 7, 2014.
- Glanz, K; Rimer, B; Viswanath, KV (editors). *Health behavior and health education: theory, research and practice* (4th ed). Companion materials. 2008. Available at: <http://www.med.upenn.edu/hbhe4/index.shtml>. Accessed January 7, 2014.
- Glasgow R. E., Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM Framework. *Am J. Pub Hlth* 1999;89(9):1922-1927.
- Rogers EM. *Diffusion of innovations*, 5th ed. New York: Free Press; 2003.
- Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. *Am J Prev Med.* 2012 Sep;43(3):337-50.

2. *Based on your experience/expertise:*

Are there specific existing or emerging methodologies (e.g., DUR, claims analysis, audits, observational studies, pre-post studies, etc) that should be considered to collect any or all of these assessment domains?

Specific methodologies, each of which has specific strengths and weaknesses, exist that can be employed to assess the effectiveness of the assessment domains. The examples provided in the question are not mutually exclusive (i.e., observational studies include DUR, claims analysis, audits, and pre-post studies, and a DUR could be conducted through chart review, a prospective observational study, or review of claims or electronic medical records). Additional methods to consider include time-series analyses, simulations and modelling approaches, pragmatic trial designs, factorial study designs, and human factor studies. The specific method should be tailored to the research objective. Table 2 summarizes various study methods, what they can measure, and associated strengths and weaknesses.

Table 2. Study Methodologies To Assess Effectiveness of REMS Programs

| Study Type | Measurement | Strengths | Disadvantages | Framework Domain(s) Study Type Can Measure |
|---------------------------|--|--|---|---|
| Secondary data collection | | | | |
| Chart review | <ul style="list-style-type: none"> ▪ Off-label use ▪ Indication for use ▪ Performance of specific clinical behaviors (e.g., patient counseling, TB or pregnancy test conducted) | Informative when detailed use information is required, for example, in-hospital administration (vs. what's in claims) | Data require extraction, Limited scope Ascertainment bias given charts not available in many settings Missing information Recording of clinical behavior in charts is incomplete, even in presence of a REMS | Reach, effectiveness, adoption, implementation, maintenance |
| Drug Utilization Study | <ul style="list-style-type: none"> ▪ Off-label use ▪ Medication use and safety event trends ▪ Clinical behaviors (e.g., patient counseling; TB or pregnancy test conducted) | Off-label use descriptive studies Can measure risk minimization with long-term data (Did exposure decrease? Did event rates decline?) | Descriptive data Requires sufficient uptake of drug before assessment can be made Trends require several years of data; trends not possible around launch Recording of clinical behavior in charts is incomplete, even in presence of a REMS | Reach, effectiveness, adoption, implementation, maintenance |

| Study Type | Measurement | Strengths | Disadvantages | Framework Domain(s) Study Type Can Measure |
|---|--|---|---|---|
| Retrospective pharmacoepi- demiology database study | Event rates Association between medication and outcome | Examines risk outcomes of interest Uses retrospective data, relatively easy to collect Could be single point in time or could conduct time series analyses to look at trends over time Relatively easy to identify several patient cohorts to serve as comparators | Requires sufficient prescription use and number of patients of interest Adjustment for confounding by indication and adjustment for disease severity may be difficult to accomplish May require validation of the study outcome Recording of some outcomes will be incomplete (e.g., pregnancies in particular) Drug sharing may not be recorded. | Effectiveness, adoption, implementation, maintenance |
| Primary data collection | | | | |
| Prospective pharmacoepi- demiology studies | Event rates Association between medication and outcome Performance of specific clinical behaviors (e.g., patient counseling; TB or pregnancy test conducted) | Examines risk outcomes of interest; Could be conducted at single point in time or implemented as a time-series analyses | Generalizability beyond sample— who participates (selection bias) Hard to interpret depending on whether a control or a comparison condition is used Not useful/applicable if event of interest is rare Adjustment for confounding by indication and by disease severity may be difficult to accomplish | Reach, effectiveness, adoption, implementation, resource utilization, maintenance |

| Study Type | Measurement | Strengths | Disadvantages | Framework Domain(s) Study Type Can Measure |
|--|--|---|---|--|
| Patient, pharmacist, or physician survey | <p>Knowledge of and beliefs about medications; behavior; resource utilization</p> <p>Outcome information can be collected via surveys.</p> | <p>Most rigorous method by which to determine what the respondent knows or believes</p> | <p>Depending on the process whereby the sample was generated, external generalizability may be an issue:</p> <ul style="list-style-type: none"> ▪ Sampling frame is typically not available since a list of the universe of patients, users, and/or prescribers may not always be readily available ▪ Interpretations often drawn from small samples that may reflect selection biases different from those of enrolled patients <p>Potential reporting bias (desirable behaviors vs. true behaviors)</p> | <p>Reach, effectiveness, adoption, implementation, maintenance; Stakeholder-centered; Resource utilization</p> |
| Process study (analysis of REMS process) | <p>Did process for disseminating information work?</p> <ul style="list-style-type: none"> ▪ Number (%) of Medication Guides or Dear Health Care Provider letters given/sent ▪ Number (%) of pharmacies/ physicians enrolled ▪ Number (%) of prescriptions in approved settings ▪ etc. <p>Sources include REMS database(s) and audits</p> | <p>Provides information on how well the process worked</p> <p>Allows for root cause analysis of the process</p> | <p>Does not provide information on outcomes (was the risk mitigated)</p> | <p>Reach, adoption, implementation, maintenance, resource utilization</p> |

| Study Type | Measurement | Strengths | Disadvantages | Framework Domain(s) Study Type Can Measure |
|---|---|--|--|---|
| Dismantling design (i.e., factorial design) | One option is a “dismantling” evaluation design (Glasgow 2005). This would involve randomly assigning different aspects of the risk minimization program to be implemented in different geographic regions. An evaluation could then be conducted in which each region is compared to the others. Viable when a risk minimization intervention or program has been designed to include different intervention components in different regions | Provides one or more comparator arms, thus increasing interpretability of the evaluation findings Helps identify the critical “active ingredient(s)” in the risk minimization intervention, thus helping to make REMS programs more efficient and less burdensome | Not always permitted or allowed by regulators | Effectiveness, maintenance |
| Pragmatic clinical or behavioral trial | Random assignment to either intervention or comparator (e.g., “usual care” or “minimal intervention”) Pretest and posttest with multiple follow-up evaluation points | Provides a comparator arm, thus increasing interpretability of the evaluation | Not always permitted/allowed by regulators Care must be taken to not affect the practice of “usual care” through study design/ procedures | Effectiveness, maintenance |

| Study Type | Measurement | Strengths | Disadvantages | Framework Domain(s) Study Type Can Measure |
|---------------------------------------|---|---|--|---|
| Quasi-experimental (no randomization) | Nonrandom assignment to intervention or random assignment to either intervention or comparator (e.g., "usual care" or "minimal intervention") Pretest and posttest with multiple follow-up evaluations Nonequivalent group design Regression discontinuity | Provides a comparator arm, thus increasing interpretability of the evaluation | Not always permitted/allowed by regulators | Effectiveness, maintenance |
| Simulation/modeling techniques | | | Limitations on data available | Resource utilization to guide estimations of model parameters |

Reference

Glasgow R. E., Magid DJ, Beck A, Ritzowoller D, Estabrooks PA. Practical clinical trials for translating research to practice: designs and measurements recommendations. *Med Care* 2005;43(6):551-557.

3. *What metric(s) (i.e., numerator, denominator, value, units) could be used as a standard for assessing reach, effectiveness, adoption, implementation and maintenance? For each recommended metric, what data sources are available to collect such information? Any commentary around this?*

The metrics used as a standard for assessing RE-AIM should be driven by the specific study objective and method used to assess the effectiveness of each domain. In addition, the extent of the overall REMS evaluation undertaken should be proportionate to risk. Each metric has relative strengths and weaknesses that should be considered. Potential metrics associated with different domains and study types are provided in Table 3.

Table 3. Metrics to Assess Domains of REMS Assessment

| Domain | Metric | Method | Data Source |
|-----------------------------------|---|--|--|
| Evidence-based program components | Extent to which program was developed using components that have been shown to be effective and/or extent to which program components have undergone piloting | Assessment of method(s) used for developing program components | Primary data collection (e.g., piloting results); Systematic reviews of relevant literature (e.g., Cochrane Collaboration, reviews undertaken by MAH); Review of "gray" literature such as that available from FDA and other regulators' websites (e.g., FDA Advisory Committee briefing materials, meeting minutes; FDA Risk Communication Advisory Committee meeting minutes and results). |

*Standardization and Evaluation of Risk Evaluation and Mitigation Strategies (REMS)—
Public Meeting on July 25-26, 2013: Response to FDA Questions*

| Domain | Metric | Method | Data Source |
|---|--|---|---|
| Stakeholder-centered program components | Extent to which REMS program components were created using input from appropriate stakeholder group(s) | Focus groups; Stakeholder panels; stakeholder advisory committees; relevant advocacy groups; stakeholder surveys or interviews. See: Principles of Community Engagement, Second Edition, 2011 (NIH Clinical Translational Science Awards) for additional methods. | Primary data collection; supplemented by review of relevant published literature if applicable. |
| Reach | Extent to which targeted audience (number and percentage) was reached and is participating in the program (e.g., patients, HCPs) and the degree to which they are representative of the target population(s) | Surveys/internal audits of REMS program records | Primary data collection |
| Effectiveness | Number and percentage who demonstrate requisite "knowledge" | Survey | Primary data collection |
| | Attitudes, behavioral intent, behaviors | Survey | Primary data collection |
| | Functional health status Quality of life Satisfaction | Survey; retrospective analysis of secondary data sources | Primary data collection; chart or electronic medical record review |
| | Incidence of outcome or relative risk | Retrospective chart or database study | Medical charts, claims data, electronic medical records |
| Adoption | Number (%) implementing desired behavior or number (%) surveyed or number (%) in database meeting criteria | Survey, database study, chart review | Primary data collection, claims data, electronic medical records, or medical charts |
| | Number of inappropriate prescriptions/number of total prescriptions | Database study | Claims or electronic medical record review |

*Standardization and Evaluation of Risk Evaluation and Mitigation Strategies (REMS)—
Public Meeting on July 25-26, 2013: Response to FDA Questions*

| Domain | Metric | Method | Data Source |
|--------------------------------|--|---|---|
| Implementation | Number of trained/number of potential HCPs in target audience | Tally | Pharmacy data |
| | Number and types of barriers to implementation identified | Survey/interview | Primary data collection |
| | Number (%) of REMS materials (e.g., Medication guides or other communication tools) distributed | Program records and/or internal audits Patient or HCP survey | Primary data collection |
| | Number (%) of REMS activities (e.g., screenings, tests) delivered in the specified amount and frequency and to the specified audience(s) | Program records and/or internal audits | Primary data collection |
| Maintenance and sustainability | Extent to which the program has been adopted at the target sites into the existing workflow process, is being delivered as planned ("process measures"), and outcome measures reflect impact | Drug utilization study across multiple years Repeated survey | Primary data collection via surveys and audits; medical charts, claims data, electronic medical records |
| Resource utilization | Number of additional staff hired to implement program Amount of time spent training staff to implement program and to monitor degree to which program is being implemented consistently Amount of extra funding needed to implement and maintain REMS program at the level of both the implementing site(s) and the central office Amount of waiting or travel time for patients Patient costs associated with REMS-required tests, visits, etc. | Survey/interview; internal audits | Primary data collection Information from databases (e.g., electronic medical records) |

4. *What data analysis and statistical methods are recommended for periodically and serially analyzing REMS performance over time? (How) can target performance thresholds be established and used to determine how well the program is meeting its goal(s)? Any commentary around this?*

Statistical Methods for Periodically and Serially Analyzing REMS Performance Over Time

One approach to serially analysing data over time is the “interrupted time series,” which has been referred to as the strongest quasi-experimental design to evaluate the impacts of health policy interventions in situations where randomization is not feasible (Zhang et al., 2011). It has been used to evaluate education and policy interventions (Biglan et al., 2000; Catalano and Serxner, 1987; Wagner et al., 2002), including tobacco control policies (Bernat et al., 2013) and the effects of FDA drug regulatory policy (Libby et al., 2007; Morgan et al., 2007; Morrato et al., 2010).

References for Time-Series Analysis

- Bernat DH, Maldonado-Molina M, Hyland A, Wagenaar AC. Effects of smoke-free laws on alcohol-related car crashes in California and New York: time series analyses from 1982 to 2008. *Am J Public Health*. 2013 Feb;103(2):214-20.
- Biglan A, Ary D, Wagenaar AC. The value of interrupted time-series experiments for community intervention research. *Prev Sci*. 2000 Mar;1(1):31-49.
- Catalano R, Serxner S. Time-series designs of potential interest to epidemiologists. *Am J Epidemiol*. 1987 Oct;126(4):724-31.
- Libby AM, Brent DA, Morrato EH, Orton HD, Allen R, Valuck RJ. Decline in treatment of pediatric depression after FDA advisory on risk of suicidality with SSRIs. *Am J Psychiatry*. 2007 Jun;164(6):884-91.
- Morgan OW, Griffiths C, Majeed A. Interrupted time-series analysis of regulations to reduce paracetamol (acetaminophen) poisoning. *PLoS Med*. 2007 Apr;4(4):e105.
- Morrato EH, Druss B, Hartung DM, Valuck RJ, Allen RR, Campagna E, et al. Metabolic testing rates in 3 state Medicaid programs after FDA warnings and ADA/APA recommendations for second-generation antipsychotic drugs. *Arch Gen Psychiatry*. 2010 Jan;67(1):17-24.
- Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther*. 2002 Aug;27(4):299-309.

Zhang F, Wagner AK, Ross-Degnan D. Simulation-based power calculation for designing interrupted time series analyses of health policy interventions. *J Clin Epidemiol*. 2011 Nov;64(11):1252-61.

Mixed Methods

Another approach is to use mixed methods, integrating qualitative methods with quantitative methods. The underlying logic of using a mixed-methods approach is that neither quantitative nor qualitative methods are sufficient alone to capture all the necessary information regarding program implementation and impact. When used in combination, both quantitative and qualitative data may yield a more complete and complementary analysis of REMS performance over time (Creswell et al., 2004).

Creswell (2012) has identified several mixed-method design types most relevant to dissemination and implementation research:

- Convergent parallel design—simultaneous data collection, merging data in the analysis
- Explanatory sequential design—qualitative data are used to explain/understand quantitative data already collected
- Exploratory sequential design—qualitative data are collected first to inform the quantitative analysis
- Embedded design—one form of data is embedded in the other, e.g., ethnographic analysis as part of an epidemiologic evaluation of REMS outcomes in a health care system
- Multiphase design—a series of phases of separate studies, e.g., sequential key informant interviews over time

References

Creswell JW. *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Boston, MA: Pearson Education Inc.; 2012.

Creswell JW, Fetters MD, Ivankova NV. Designing a mixed methods study in primary care. *Ann Fam Med*. 2004 Jan-Feb;2(1):7-12.

Statistical Process Control Charts

A third method to consider is data displayed as statistical process control charts. This is a statistical method for quality control measurement used in industry and hospital epidemiology, typically under the heading of total quality and Six Sigma. Control charts are also increasingly being used in public health surveillance and health care (Hanslik et al., 2001; Woodall, 2006). They aid in understanding processes and evaluating the

impact of changes and whether changes lead to improvement, deterioration, or status quo in quality metrics. Methods have been described thoroughly (Woodall, 2011), and the choice of which control chart to use depends on the type of data to be plotted.

References

- Hanslik T, Boelle PY, Flahault A. The control chart: an epidemiological tool for public health monitoring. *Public Health*. 2001 Jul;115(4):277-81.
- Woodall WH. The use of control charts in health-care and public-health surveillance. *J Quality Technol*. April 2006;38(2):89-104.
- Woodall WH. The use of control charts in healthcare (chapter 13). In: Faltin F, Kenett R, Ruggeri F (editors). *Statistical methods in healthcare*. Wiley; 2011.

Social Media Data Analysis

Another emerging method that may be relevant for some REMS programs is in the area of social media data analysis (Chary et al., 2013; Tufekci, 2013).

References

- Chary M, Genes N, McKenzie A, Manini AF. Leveraging social networks for toxicovigilance. *J Med Toxicol*. 2013 Jun;9(2):184-91.
- Tufekci Z. Big data: pitfalls, methods and concepts for an emergent field. 07 March 2013. Available at: <http://ssrn.com/abstract=2229952>. Accessed January 7, 2014.

Target Thresholds

All REMS programs should have *a priori* established target performance thresholds. Different REMS programs may employ different threshold levels for the same or similar process and outcome measures due to a variety of factors (e.g., different patient or HCP populations being targeted or different geographic areas covered). The same thresholds should not be assumed to be equally relevant across different REMS programs. Factors to consider include (1) ideal or desired performance (e.g., correct knowledge level, percentage compliance), (2) the extent to which a REMS program can influence the intended audience (e.g., a REMS program may educate a physician on appropriate prescribing and the prescriber may be informed of the potential risk, yet still choose to prescribe the medication to a patient in a manner that is inconsistent with appropriate prescribing practices), and (3) the specific unit of measurement/analysis.

Question C

1. *Does the working group have any additional recommendations for submission to the docket?*

C1. Comparison of FDA and EMA Approaches to Risk Management

A comparison of risk management regulations between the FDA and the European Medicines Agency (EMA) suggests that the current FDA guidance (FDA, 2009) and EMA guidance (EMA, 2012) documents share similar objectives with regard to the identification, monitoring, and minimization of risk and, as a consequence, have similar data requirements (Lis et al., 2011; Lis et al., 2012). However, one notable difference between the two approaches is that EMA Risk Management Module V makes a distinction between “routine” versus “additional” risk minimization measures (aRMM). Routine risk minimization activities refer to product labeling (e.g., Summary of Product Characteristics and Patient Information Leaflet), limitations on the number of units that can be prescribed in any single prescription (“drug pack size”), and the legal status of the product. Collectively, the routine activities represent the “foundation” upon which “additional” risk minimization measures are determined. Additional risk minimization measures include activities to address safe and appropriate use of a product, including, for example, educational interventions, clinical decision aides, reminder systems, restricted access or limited distribution programs, package design elements, and special screening or laboratory tests.

The two regulatory agencies are also aligned in the emphasis on communicating risk and/or benefit-risk to patients. In the United States (US), Medication Guides are the designated vehicle for communicating risk information to patients; whereas in the European Union (EU), the Patient Information Leaflet is used. In terms of HCP communication, the Package Insert, Dear Health Care Professional Letters, brochures, and other forms of messaging are all permitted for use as part of the REMS program component Health Care Communication Plan. In the EU, the Summary of Product Characteristics is the primary tool used to communicate benefit and risk to health care professionals.

In the US, assessment of patient access and burden imposed by the risk minimization program is required to be addressed, specifically in regard to programs involving elements to assure safe use (FDA, 2009, line 756). In actual practice, however, assessment of access and burden imposed by REMS has not been consistently or comprehensively assessed. One recent exception has been the Class REMS for Extended-Release and Long-Acting Opioids, in which an eighth assessment measure was included to measure access to analgesic treatment.

In the EU, the EMA's Guideline on Good Pharmacovigilance Practices (GVP) Module XVI specifies that all evaluations should assess both implementation fidelity and program impact or effectiveness via the use of process and outcome indicators, respectively (EMA, 2013). The CIOMS IX Working Group proposes that evaluations utilize a framework that incorporates domains such as the five included in RE-AIM: reach, adoption, implementation, effectiveness, and maintenance (CIOMS, 2013). In addition, CIOMS IX also notes that program implementation (in terms of content and frequency and dosage amount delivered) should be assessed at frequent intervals so that corrective action can be expedited in "real time" (e.g., during implementation) as part of a continuous quality improvement cycle.

The CIOMS IX Working Group also emphasizes three additional points in regard to evaluation design:

- *Barriers and facilitators to successful program implementation.* The importance of capturing both contextual and moderating factors (e.g., degree of staff training, the existence of a local risk minimization program champion) that may serve to either impede or facilitate program implementation in different areas, which can, in turn, lead to variations in outcomes across different locales.
- *Determination of thresholds of program "success."* The importance of prespecifying thresholds for determining program "success" or effectiveness cannot be underestimated. To guide this process, CIOMS IX recommends that program planners take three factors into account in estimating the magnitude of the expected effect and therefore the choice of a threshold for determining the success of a risk minimization program:
 - (1) The impact of the risk (i.e., likelihood × severity of harm)
 - (2) The desired level of the risk minimization effort in relation to the benefits provided by the drug, the indication, and the vulnerability of the patients prescribed this drug (e.g., it is reasonable to expect a stricter measure of performance for a medicinal product used in children in the prevention of a disease compared with the same medicinal product used as a therapeutic option in adults)
 - (3) What is practical and feasible given resources and/or time. For some REMS programs that are imposed as a condition of marketing approval, a success threshold can potentially be defined using relevant data from phase 3 trials.
- *Selection of first evaluation assessment and periodicity of subsequent assessments.* In regard to the timing of evaluation assessments, the first time point to be selected for evaluation should avoid conflating program implementation with program outcomes, and any time point falling prior to the renewal of a marketing authorization should be selected such that results on the

evaluation of effectiveness will be made available for the coming regulatory review.

For globally marketed products, both EMA Module XVI and the CIOMS document stress the importance of developing risk minimization programs and evaluation plans that can be implemented across multiple geographic regions and markets. Novel evaluation methods are in the process of being developed and piloted. It would be helpful to consider harmonizing evaluation measures of initiatives to mitigate risks in the US and EU for the same product so that comparative assessments of prevention efforts across regions can be conducted.

C2. Balancing Knowledge Gained vs. Burden on Health Care Systems

In the US, REMS programs that have elements to assure safe use have in some circumstances required active involvement of pharmacies (utilizing the systematic functionality of e-pharmacy management systems) to verify prescriber certification (e.g., the shared REMS for transmucosal immediate-release fentanyl products). The feasibility and burden of this approach could be assessed. In terms of assessing the impact and effectiveness of REMS programs, such an assessment not only provides insight into the impact and value of the program, but can be used to calculate the cost-effectiveness and population attributable fraction associated with the REMS intervention(s) against the savings incurred through compliance/adherence to the medication guidelines (and off-set against burden). To derive this information, electronic medical records (EMR) and registries or EMR-enhanced administrative data such as those compiled by the Mini-Sentinel project could be used. Additional insight might be gained from surveys to patients who discontinue the medication.

C3. Measuring the Impact of REMS on Patient Access to Needed Medications

As FDA has stated, REMS programs are intended to improve the framework for benefit-risk balance of medicinal products. However, this goal may be undermined due to unintended or undue restriction of patient access to the products. Possible ways to measure the impact of a REMS program on patient access are inclusion of questions in prescriber and patient surveys to measure access, assess the type and degree of perceived barriers to access, and ask patients and prescribers about the current level of access to the specified drug product

References

CIOMS (Council for International Organizations of Medical Sciences). CIOMS Working Group IX on "Practical Considerations for Development and Application of a Toolkit for Medicinal Product Risk Management." 2013. Available at: <http://www.cioms.ch/index.php/working-group-ix>. Accessed January 7, 2014.

European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module V – Risk management systems. June 22, 2012. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/06/WC500129134.pdf. Accessed January 7, 2014.

European Medicines Agency. Guideline on good pharmacovigilance practices (GVP). Module XVI– Risk minimisation measures: selection of tools and effectiveness indicators (draft guidance). June 6, 2013. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/06/WC500144010.pdf. Accessed January 7, 2014.

Food and Drug Administration. Guidance for industry. Format and content of proposed risk evaluation and mitigation strategies (REMS), REMS assessments, and proposed REMS modifications (draft guidance). US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER); September 2009. Available at: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM184128.pdf>. Accessed January 7, 2014.

Lis Y, Guo JJ, Roberts MH, et al. A comparison of US Food and Drug Administration and European Medicines Agency regulations for pharmaceutical risk management: report of the International Society for Pharmacoeconomics and Outcomes Research Risk Management Working Group. *ISPOR Connections*. 2011;17:10-13. Available at: <http://www.ispor.org/sigs/riskbenefitmanagement/A-comparison-of-US-Food-and-Drug-Administration-and-European-Medicines-Agency.pdf>. Accessed January 7, 2014.

Lis Y, Roberts MH, Kamble S, Guo JJ, Raisch DW. Comparisons of Food and Drug Administration and European Medicines Agency risk management implementation for recent pharmaceutical approvals: report of the International Society for Pharmacoeconomics and Outcomes Research Risk Benefit Management Working Group. *Value Health*. 2012 Dec;15(8):1108-18.

Acknowledgement

The authors would like to thank all those ISPE members that provided comments and critical review to this response including, but not limited to, Kenneth Hornbuckle, Douglas Watson, Elizabeth Andrews, Nancy Dreyer and Katarzyna Sablinska.