12-13-2012

Building the Informatics Infrastructure for Comparative Effectiveness Research (CER): A Review of the Grey Literature

Marianne Hamilton Lopez
AcademyHealth, marianne.hamiltonlopez@academyhealth.org

Rebecca Singer Cohen
AcademyHealth, rebecca.singercohen@academyhealth.org

Erin Holve
AcademyHealth, erin.holve@academyhealth.org

Recommended Citation
Hamilton Lopez M., Singer Cohen R. and Holve E., "Building the Informatics Infrastructure for Comparative Effectiveness Research (CER): A Review of the Grey Literature," EDM Forum, AcademyHealth, December 2012.

EDM Forum

Issue Brief

Building the Informatics Infrastructure for Comparative Effectiveness Research (CER): A Review of the Grey Literature

About the EDM Forum

The Electronic Data Methods (EDM) Forum is a three-year grant from the Agency for Healthcare Research and Quality (AHRQ) to advance the national dialogue on the use of electronic clinical data for the conduct of comparative effectiveness research (CER), patient-centered outcomes research (PCOR), and quality improvement (QI). The EDM Forum facilitates exchange and collaboration between eleven AHRQ-funded projects, including: the Prospective Outcome Systems using Patient-specific Electronic data to Compare Tests and therapies (PROSPECT) studies; the Scalable Distributed Research Networks for CER, the Enhanced Registries for QI and CER, as well as other relevant health IT initiatives. The EDM Forum and the research projects connected to the Forum are funded through the American Recovery and Reinvestment Act of 2009 (ARRA).

AcademyHealth

1. Introduction

In 2011, the Electronic Data Methods (EDM) Forum conducted a search of the peer-reviewed literature at the intersection of CER and clinical informatics. A three-step process was utilized, including a structured search of PubMed, manual reviews of articles from selected publication lists, and manual reviews of papers on the use of electronic clinical data (ECD) for CER.[1] Based on the review, investigators determined that the peer-reviewed literature on CER leveraging ECD is still in the early stages of development, and does not cover some topics of great interest to the community (e.g., single point access).

Upon reviewing the findings from the analysis of peer-reviewed literature, the EDM Forum Steering Committee noted that it is possible that emerging work in this area may not have reached peer-review status or that because many studies in this space are supported by government entities or non-governmental organizations, there are alternative publication strategies for this work that exist outside of traditional peer-review channels. As a result, the EDM Forum Steering Committee suggested a structured review of the grey – or non-commercial — literature[2] be conducted to under-

About the EDM Forum Peer-Reviewed Literature Review:

AcademyHealth researchers identified one hundred thirty-two peer-reviewed articles relevant to CER and electronic clinical data. Of these, 88 articles were selected for analysis in a recently published manuscript in Medical Care. Three types of articles were identified, including papers that: (1) provide historical context or frameworks for using clinical informatics for research; (2) describe platforms and projects, and (3) discuss issues, challenges and applications of natural language processing (NLP). Two cross-cutting themes emerged: the challenges of conducting research in the absence of standardized ontologies and data collection; and unique data governance concerns related to the transfer, storage, de-identification, and access to ECD. The authors identified several current gaps on important topics such as the use of clinical informatics for cohort identification, cloud computing, and single point access to research data. For more information, please see *Hamilton Lopez M., et al. Building the Informatics Infrastructure for Comparative Effectiveness Research (CER): A Review of the Literature. Med Care. 2012 Jul;50 Suppl:S38-48.*

The EDM Forum is supported by the Agency for Healthcare Research and Quality (AHRQ) through the American Recovery & Reinvestment Act of 2009, Grant U13 HS19564-01.

Figure 1: Cross-Cutting Keywords Relevant to CER and ECD

A list of cross-cutting keywords was developed for the peer-reviewed literature review. In the grey literature review, a subset of the keywords (see below) was used in a series of searches.

"Cloud Computing"	"Metadata"
"Cohort Identification"	"Natural Language Processing"
"Data Use and Quality"	"Patient Involvement"
"De-Identification"	"Research Networks"
"Governance"	"Security"
"Institutional Review Board"	"Single Point Access"
"Learning Healthcare system"	"Standardized Data Collection"
"Library of Phenotypes"	

stand the emerging literature on developing the CER infrastructure with ECD.

Prior to this review, there has been no existing structured search of CER using ECD in the grey literature. To address this gap, our efforts included developing a structured methodology; estimating the total volume of grey literature on ECD used for research and quality improvement (QI); and addressing the Steering Committee's questions about whether the peer-reviewed search may have systematically excluded major perspectives that might be found in the grey literature. The grey literature resulting from the search has been characterized by its size, area of focus, and gaps.

2. Methodology

2.1 Study Design and Search Strategy

As expert panels have noted, at present there are no accepted standards for conducting systematic reviews of the grey literature.[3, 4, 5] To search the grey literature most effectively[6], the EDM Forum staff employed three separate search strategies to ensure a sufficiently robust search:

2.1.1 Structured searches for key concepts in grey literature-producing organizations.
The list of 47 grey literature-producing organizations (see Appendix A) that were reviewed was developed based on three major sources:

1. A subset of sources that focus on CER or information technology (IT) from the Master List of Grey Literature-Producing Organizations developed by AcademyHealth for the National Library of Medicine;[7]
2. Organizations identified by the EDM Forum Steering Committee and staff; and
3. Input from EDM Forum consultants and experts in the fields of health services research, library science, and clinical informatics.[8]

The initial search string used was "comparative effectiveness." If there were more than 50 results, then subsequent search strings were employed, including ("comparative effectiveness" AND informatics) and ("comparative effectiveness" AND "health information technology"). EDM Forum staff manually searched each organization's website for sections of the site focused on Health IT or CER, as well as the publications page for relevant documents or web pages. If no such subject headings were available, a general site-wide search was conducted.

2.1.2 Structured Google Searches.
This effort focused on a set of keyword (KW) and KW search strings of terms known to be associated with key projects and programs relevant to the EDM Forum and its affiliated research teams. Using the Google search engine, the following searches were conducted:

1. "comparative effectiveness"
2. "comparative effectiveness" AND "informatics"
3. "comparative effectiveness" AND "health information technology"

A similar process was followed for a set of additional cross-cutting KWs developed for the peer-reviewed literature search (figure 1). The KWs were searched independently and then as part of structured search strings:

1. "KW"
2. "KW" AND "comparative effectiveness"
3. "KW" AND "health information technology"

Based on emerging best practices from pre-existing grey literature systematic reviews, EDM Forum staff selected for analysis the first 50 Google results per search.[9,10] This decision is consistent with Google's algorithm to surface the most relevant results.

2.1.3 Manual structured searches of websites produced by known CER projects and platforms relevant to the EDM Forum.
For both the peer-reviewed and grey literature searches a list of appropriate websites from relevant CER projects and platforms using ECD was developed and searched based on input and discussions with experts working on CER and informatics, including members of the EDM Forum Steering Committee. These projects and platforms included: caBIG (cancer Biomedical Informatics Grid), DARTNet, DEcIDE (Developing Evidence to Inform Decisions about Effectiveness), HMORN (HMO Research Network), iDASH (integrating data for analysis, anonymization, and sharing), i2b2 (Informatics for Integrating Biology and the Bedside), OMOP (Observational Medical Outcomes Partnership), PhysioMIMI (Multi-Modality, Multi-Resource Information Integration environment), REDCap (Research Electronic Data Capture), Sentinel Initiative and Mini-Sentinel, SHARP Program (Strategic Health IT Advanced Research Projects), TRIAD (OSU Clinical and Translational Science Awards), and VINCI (VA Informatics and

Computing Infrastructure). Additionally, a Google search was conducted for each CER project or platform to make sure that the website search was surfacing every relevant result.

2.2 Search Criteria and Strategy

Grey literature documents and Web pages were *included* in the analysis if they were:

- relevant to using clinical informatics and ECD to conduct CER, and
- had *not* been published in a peer-reviewed journal (i.e., published by a commercial publisher).

In cases where multiple articles or reports on the same topic were published by the same author, the latter, more up-to-date article was included.

Documents and Web pages were *excluded* if they were:

- not explicitly related to clinical informatics and ECD and/or CER;
- focused more on the clinical aspect of a study rather than the informatics aspect; and
- focused on genetic rather than clinical or translational research.

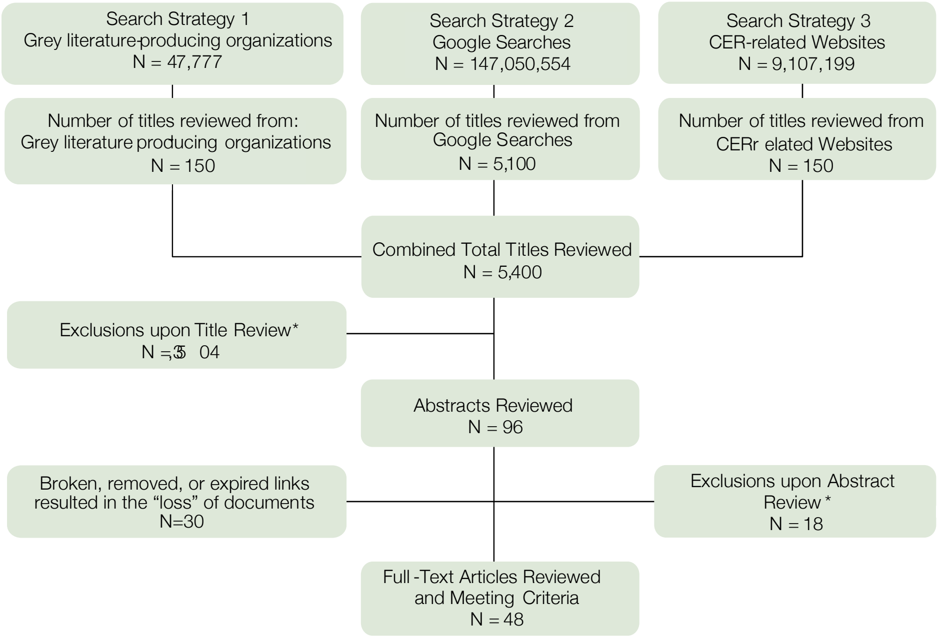
Two AcademyHealth staff conducted the inclusion/exclusion review process. Inconsistencies between the reviewers' judgments regarding inclusion or exclusion of articles were deliberated and resolved.

A total of 78 documents and/or web pages met the criteria for initial full-review. However, a significant portion of the links for these documents were broken, removed, or expired. Excluding these sources resulted in the "loss" of 30 documents and/or Web pages because they were no longer available online.

Forty-eight grey literature documents and/or Web pages (see Appendix B) met the criteria for inclusion analysis and were abstracted into an abstraction form developed by the EDM Forum (Appendix C).

An AcademyHealth reviewer abstracted citation information to the extent available. Where no suggested citation existed, the closest Vancouver style citation that could be created was developed. Sections such as a stated methodology section, funder, or suggested citation, vocabulary, themes, and results were also abstracted.

Figure 2: Flow Diagram of Structured Search Results



**Reviewers applied the inclusion/exclusion criteria described in the Methods section.*

3. Findings

Based on a review of the relevant literature, a total estimated volume of grey literature emerges, as do a set of emerging themes in the grey literature, which may be compared to the peer-reviewed search. Both issues are discussed in the following section.

3.1. Volume of Grey Literature

The three search strategies yielded more than 156 million potential documents and Web pages. Of these, 5,400 titles were reviewed. Only 48 documents and Web pages were ultimately identified as being most relevant for this search. See figure 2.

3.2. Characteristics of the Grey Literature: Themes in the grey literature compared to the peer-reviewed search

The list of keywords from the peer-reviewed search was used to characterize areas of focus. After the records were retrieved, each document and/or Web page was reviewed and coded with a primary keyword[11] (see figure 1) in order to assess the extent to which specific topics are currently addressed in the grey literature.

There was good parity between some of the topics. A comparable percentage of the peer-reviewed and grey literature documents and/or Web pages were coded as focusing on "research networks" and "standardized data collection." And the peer-reviewed and grey literature reviews shared three similar gaps in the literature; no literature in either search was identified as focusing on "single point access" to research data, "cloud computing" or "cohort identification[12]." However, there were differences between the results of the two searches, including:

- The identified primary code "Natural Language Processing" (NLP) was less prevalent as a proportion of the grey literature (0 percent) compared to the peer-reviewed literature (23 percent);
- The identified primary code "Security"

Figure 3: Comparison of Grey and Peer-Reviewed Code Results

Codes	Grey Literature (n=48)	Peer Reviewed Literature (n=132)
The Learning Health Care System *	29%	1%
Data Use and Quality	21%	13%
Security*	13%	3%
Overview	13%	17%
Research Networks	8%	7%
Identifiers and De-Identification	6%	3%
Standardized Data Collection	4%	5%
Metadata	4%	2%
Patient Involvement	2%	2%
Natural Language Processing (NLP)*	0%	23%
Institutional Review Boards (IRB)	0%	2%
Governance	0%	1%
Library of Phenotypes	0%	1%
Single Point Access	0%	0%
Cloud Computing	0%	0%
Cohort Identification	0%	0%

Statistically significant difference observed between the proportional search of grey literature and peer reviewed literature on specific topics.

was more prevalent as a proportion of the grey literature (13 percent) compared to the peer-reviewed literature (3 percent); and

- The identified primary code “Learning Health Care System” was more prevalent as a proportion of the grey literature (29 percent) compared to the peer-reviewed literature (1 percent).

4. Discussion

The goal of this review was to develop and apply a structured methodology to estimate the volume of grey literature at the intersection of ECD and CER and to characterize the results. Due to the nascent nature of this area of study (and the absence of a standardized terminology or taxonomy), both tasks proved complex. Early lessons learned from the search suggest that future examinations of the grey literature will need to account for identification and quality of the grey literature. The ability to estimate the volume of the grey literature was impacted by the number of documents and Web pages identified by the search results. Combined, the initial searches yielded over 156 million potentially relevant documents and web pages. The volume of the results from Google searches was not a surprise; however, the high number of search results generated by searches of the specialized websites was unexpected because these sites are considerably more specific in their scope. Perhaps more importantly, reviewers were struck by the low number of relevant sources in proportion to the extremely large results produced by the various searches. A total of 5,400 titles were reviewed - the top 50 results from each of the site searches (see figure 2), of which less than 1 percent were relevant for the EDM Forum. This low hit rate suggests that the semantic search strategies employed through the Google algorithm may not have been well suited for the purposes of this search.

The absence of a structured vocabulary within this emerging area of study combined with the ever-expanding nature of the Internet contributed to issues of classification. The list of cross-cutting keywords developed for the peer-reviewed literature search was consistent with terms found in the research and informatics communities. Further research is needed to better understand how these same concepts are discussed and defined by representatives of other communities such as government, industry or patients/consumers. As a result, it may be important to review and refine our current search strategy and employ less sensitive searches to better capture various terminologies.

Based on this review’s findings, a consideration for the future is that — unlike the peer-reviewed findings — the material produced outside of commercial publications rarely contains a methods section. Furthermore, the grey literature does not present material in a standard format that would allow the reader to evaluate the products’ rigor and usefulness. By and large, the length and structure of grey documents appear to be determined by the amount and depth of material available and the perceived interest and attention-span of the audience.

For example, of the 48 documents and Web pages included in the final analysis:

- Ten percent included a recommended citation and only 15 percent contained an explicit section describing the methods by which data or information was collected.
- Twenty-seven percent explicitly disclosed the funding mechanism or agencies that supported the research or publication.
- Sixty-nine percent contained citations or references.
- Eighty-three percent contained a date of publication (2004- 2011). Of those documents that did include a date, 43 percent were published in 2010, and 25 percent were published in 2009.

It was interesting to find that a number of cross-cutting concepts relevant to CER and ECD, and identified for the peer-reviewed literature, were addressed in the grey literature. This suggests that discussions are underway in a number of communities, about the challenges and benefits of leveraging ECD for research and QI.

At the same time, there appears to be three major topical differences in the focus of the grey and peer-reviewed literature related to technical topics such as security,

Limitations to searching the grey literature:

- *Semantic searches may miss important links or files that do not include key terms in the title* (which may not always accurately describe the content of a document), or other metadata associated with a documents. Searches using each of the websites' own search toolbars (on rare occasions powered by Google) were sometimes impacted by text limits in the toolbar, which made complex semantic searches difficult.
- *Google uses an algorithm which factors in a user's previous searches and selected or clicked results to "learn" individual preferences, which could bias or even negatively impact search results.* There is also a possibility that expanding the number of results included for consideration (beyond the first 50 Google results), turning off the Google algorithm, or searching other search engines, such as Bing and Google Scholar, would produce different findings. Going forward, a comparative test of the search engines using a small sample set could inform future directions of the literature review.
- *Members-only websites or password-protected sections limited access to some sources of grey literature.* Broken, moved, removed, and expired links also limited access. The transience of available material is impacted by the lack of formal indexing and standards for archiving grey literature sources (the purpose of which is served by vehicles such as PubMed for the peer-reviewed literature).[13]

policy-relevant documents, and NLP. For example, the topic of "security" was more prevalent in the grey literature perhaps because technical communities see less need for peer-reviewed publication. Additionally, likely as a result of the types of audiences that grey literature sources draw, the results from this grey literature focused heavily on current events and/or policymaking activities. Every publication dated 2009 and onward referenced recent legislation such as the American Recovery and Reinvestment (ARRA) Act, the Health Information Technology for Economic and Clinical Health Act (HITECH Act), or the Patient Protection and Affordable Care Act and their impact on targeted audience or constituencies. Policy topics, such as building a learning healthcare system are timely and broader in scope and policy audiences are more interested in timely publication and dissemination. By comparison, "natural language processing," which is used to extract relevant data from the free-text embedded in electronic health records and text documents, was absent from the grey literature search results. This is likely due to the fact that the highly technical and empirical nature of NLP validation studies is arguably better suited for traditional peer-reviewed publication.

5. Conclusion and Next Steps

Research supported by ARRA, the HITECH Act, and the Patient-Centered Outcomes Research Institute (PCORI) will likely result in new peer-reviewed and grey literature on research as well as policy issues such as research governance, and implementation strategies. The EDM Forum is committed to developing methodology to develop more systematic search strategies to bring diverse literatures together. Finding rigorous approaches to search the grey literature will be important to inform the research and policy communities about multidisciplinary viewpoints, ongoing initiatives, challenges, and proposed solutions to leverage ECD for CER, PCOR, and QI in order to improve patient outcomes.

The EDM Forum's review of the grey literature on ECD for CER was a pilot study to understand whether a search of the grey literature would add key research or policy analysis to our peer reviewed findings. As a result of this process, issues relating to identification and quality in the grey literature impacted our ability to identify and characterize the literature. We also discovered that a number of the cross-cutting keywords relevant to CER and ECD were being discussed in varying degrees in both the peer-reviewed and grey literature.

The EDM Forum team plans to continue structured searches of the grey literature as a complement to the peer-reviewed search to identify ongoing initiatives and stakeholders using ECD for improving patient outcomes. In addition, we hope to share lessons learned about ways the research can improve the utility of the grey literature to advance scientific discussion about ECD.

5.1 Refining Search Strategies for Grey Literature

In the future, the EDM Forum team may refine the search strategy to address some of the limitations of the current search such as the difficulties related to applying structured search strings within websites' search toolbars and accessing and archiving restricted, broken, moved, removed, or expired links . Future grey literature reviews will be expanded to include additional search engines, such as Bing and Google Scholar.

5.2 Providing Guidelines to Improve the Transparency and Utility of the Grey Literature

As the grey literature on ECD expands, there is an important opportunity to think more deeply about the desirable characteristics of grey literature that would make non-commercial publishing more valuable. If we believe an open marketplace of ideas should make good scholarship accessible, and that methods should be transparent, the lack of a structured methods section is an obvious problem. Future work will explore using lessons learned to provide potential guidelines to improve transparency. Also, to ensure our future searches focus on the most rigorous grey literature, forthcoming reviews may focus exclusively on abstracting and analyzing documents and Web pages that provide the project's methods.

About AcademyHealth

AcademyHealth is a leading national organization serving the fields of health services and policy research and the professionals who produce and use this important work. Together with our members, we offer programs and services that support the development and use of rigorous, relevant and timely evidence to increase the quality, accessibility, and value of health care, to reduce disparities, and to improve health. A trusted broker of information, AcademyHealth brings stakeholders together to address the current and future needs of an evolving health system, inform health policy, and translate evidence into action.

About the Authors

Marianne Hamilton Lopez is a senior manager at AcademyHealth. She can be reached at marianne.hamiltonlopez@academyhealth.org. Rebecca Singer Cohen is a senior associate and Erin Holve is a director at AcademyHealth.

Acknowledgements

AcademyHealth acknowledges the Agency for Healthcare Research and Quality (AHRQ) for its support of this work. The EDM Forum is supported by AHRQ through the American Recovery & Reinvestment Act of 2009, Grant U13 HS19564-01. AHRQ's mission is to improve the quality, safety, efficiency, and effectiveness of health care for all Americans. As 1 of 12 agencies within the Department of Health and Human Services, AHRQ supports research that helps people make more informed decisions and improves the quality of health care services. For more information, visit www.ahrq.gov.

Suggested Citation

Hamilton Lopez M., Singer Cohen R. and Holve E., "Building the Informatics Infrastructure for Comparative Effectiveness Research (CER): A Review of the Grey Literature," EDM Forum, AcademyHealth, December 2012.

Appendix A: The list of 47 reviewed grey literature-producing organizations

1. AcademyHealth
2. Agency for Healthcare Research and Quality
3. America's Health Insurance Plans
4. Alliance for Health Reform
5. American Enterprise Institute
6. American Health Information Management Association
7. American Hospital Association
8. American Medical Informatics Association
9. Avalere Health
10. Brookings Institution
11. Center for American Progress
12. Center for Health Care Strategies, Inc.
13. Commonwealth Fund
14. Connecting for Health/Markle Foundation
15. Deloitte
16. Department of Health and Human Services
17. Department of Health and Human Services/Assistant Secretary for Planning and Evaluation
18. eHealth Initiative
19. Families USA
20. Galen Institute
21. The George Washington University School of Public Health and Health Services
22. Georgetown University Health Policy Institute
23. Government Printing Office
24. Harvard School of Public Health
25. Henry J. Kaiser Family Foundation
26. Heritage Foundation
27. Institute for Healthcare Improvement
28. Institute for Policy Innovation
29. Institute of Medicine
30. Kaiser Permanente Institute for Health Policy
31. Lewin Group
32. Mathematica Policy Research
33. MedPAC
34. National Association of County and City Health Officials
35. National Association of Public Hospitals and Health Systems
36. National Business Group on Health
37. National Center for Policy Analysis
38. National Coalition on Health Care
39. National Committee for Quality Assurance
40. National Institute for Health Care Management
41. National Institutes of Health
42. National Quality Forum
43. New York Academy of Medicine
44. RAND
45. Robert Wood Johnson Foundation
46. UCLA Center for Health Policy Research
47. Urban Institute

Appendix B: Summary Table of Grey Literature Results

Source	Title	Type	Description of the grey literature product
Agency for Healthcare Research and Quality (AHRQ): Brown J, Holmes J, Syat B, et al. - Prepared by the DEcIDE Centers at the HMO Research Network and the University of Pennsylvania (June 2010)	Proof-of-principle evaluation of a distributed research network	Effective Health Care Research Report	The report focused on the distributed research network demonstration, which was designed to illustrate the functions of the central portal, the ability to conduct menu-driven distributed queries, and secure distribution and remote execution of SAS code and aggregation of the results. Lessons learned include the benefits of an incremental approach to software development and network implementation. Users preferred a “publish and subscribe” model that allows for local control of patient data while entering queries into a central portal. Security and autonomy features should be amped up with these additional functionalities.
Agency for Healthcare Research and Quality (AHRQ): Brown J, Holmes J, Maro J, et al. - Prepared by the DEcIDE Centers at the HMO Research Network Center for Education and Research on Therapeutics and the University of Pennsylvania (July 2009)	Design specifications for network prototype and cooperative to conduct population-based studies and safety surveillance.	Effective Health Care Research Report	According to the report, a distributed data network can support observational studies and allow for the prospective collection of patient data.Several design considerations must take place when building this infrastructure, including: scalability, transparency, autonomy, data heterogeneity, security, sustainability, and parsimony. Privacy considerations must be taken into account when constructing a network design. This includes being compliant with HIPAA laws as well as conducting research in a manner acceptable to the IRB. The authors propose a distributed data model in which there is a central portal controlling activities of operations/workflow and privacy, and distributed data marts, which would allow patient data to be stored in their primary location.
Agency for Healthcare Research and Quality (AHRQ): Wilson Pace, David R. West, Robert J. Valuck, Maribel Cifuentes, Elizabeth W. Staton (July 28, 2009)	Distributed Ambulatory Research in Therapeutics Network (DARTNet).	Summary Report	According to the report, the primary aim of DARTNet is to conduct observational comparative effectiveness research, and the secondary aim is to conduct a full spectrum of practice-based research. DARTNet was evaluated on 3 factors: data integrity, software functionality, and system security. The study showed that DARTNet was effective in identifying diabetic patients and conducting comparative studies on treatments and outcomes. Lessons learned included the possibility of overloading the DARTNet system and breaching security of patient data. There was also incomplete and variable patient data within the EHR, making data quality on the practice end an area for improvement.
Alliance for Health Reform (October 2008)	Health Information Technology: More Than Money	Issue Brief	The issue brief states that HIT can help promote quality, including error reduction, improved access to timely information, and patient access to better treatment information (thus allowing consumers to become more actively involved in their health care). There is no lack of agreement on the quality benefits that might derive from widespread adoption of HIT. But the challenges to get there are many. Public support is strong for information technology as a tool to improve care quality. Analysts suggest several roles the federal government can play in advancing HIT, such as providing leadership, setting standards, consumer protection and privacy policies, and offering financial incentives. The question of how best to use financial incentives to drive broader scale HIT adoption is still open.
American Academy of Orthopaedic Surgeons (AAOS): Rosemarie Nelson (March 7, 2012)	Managing Patient Data Security.	Online Article	According to this online article, HIPAA requires all healthcare covered entities (CEs)—and that includes orthopedic surgeons and their business associates (BAs)—to safeguard the privacy of patient health information. The HIPAA law also requires CEs and BAs to implement required security measures to protect patient health information. Next steps should be to perform an assessment, establish a baseline scorecard, and track compliance progress. New penalties for violating HIPAA and HITECH Act security regulations are enormous. CEs and BAs face up to $1.5 million in fines for multiple violations of a single requirement in a calendar year, as well as untold damage to reputations.
American Enterprise Institute (AEI): Stephen T. Parente (December 8, 2010)	Harnessing Health Information in Real Time: Back to the Future for a More Practical and Effective Infrastructure.	White Paper	The author offers a market-oriented alternative to the current centralized health-IT procurement approach, which he states is supplemented with insufficient bribes and penalties to achieve private-sector compliance with interoperability standards. He concludes that this more practical “back to the future” path to harnessing health information in real time can deliver long-overdue dividends in medical-fraud control, insurance-coverage administration and underwriting, and improved health care quality.

Source	Title	Type	Description of the grey literature product
American Medical Association (AMA) AmericanMedNews: Pamela Lewis Dolan (March 21, 2011)	Carelessness behind many health data breaches.	Online Article	According to this online article, for all the high-tech security work that physician practices do so that no outsiders get unauthorized access to their patients' data, one very low-tech cause of data insecurity often is overlooked: plain old forgetfulness. Policies should acknowledge that sometimes physicians or other employees need to access data at home. But practices must identify what can go wrong and look at ways of reducing the chances of those things happening. When those policies are put in place, organizations need to revisit and update their policies and monitor their systems for possible risks.
Brookings Institution: Darrell M. West (January 28, 2011)	Enabling Personalized Medicine through Health Information Technology: Advancing the Integration of Information.	White Paper	The author states that there are several ways in which personalized medicine can be enabled: (1) "meaningful use" requirements promulgated by the executive branch, (2) change driven by consumer demand for personalized medicine, (3) pilot and demonstrations projects supported by the Centers for Medicare and Medicaid Services (CMS) Innovation Center, and (4) academic-industry collaborations encouraged by the government through investment. Three "revolutions" have had a significant recent impact on health care: the medical delivery revolution involving new actors and relationships, the digital revolution and ways to convert data into knowledge, and the genomic revolution. There are several policy challenges that need to be solved: better data sharing networks, improved semantics and data coding, more balanced privacy rules, privacy and access controls, harmonization of state laws, the Catch-22 of reimbursements, rapid learning feedback mechanisms in clinical care, and predictive modeling in physician practices.
Center for American Progress: Karen Davenport (May 2007)	Navigating American Health Care: How Information Technology Can Foster Health Care Improvement.	White Paper	This paper states that health information technology can provide the information infrastructure and electronic communication systems for health care processes and quality improvement that will in turn revolutionize the American health care system. The paper also outlines immediate, medium-term, and long-term steps needed to be taken to convert the health care system into a results-based industry.
Center for Democracy & Technology: (June 2009)	Encouraging the Use of, and Rethinking Protections for De-Identified (and 'Anonymized') Health Data.	White Paper	HIPAA regulates the use, access, and disclosure of documents with patient data that is fully identifiable as "protected health information" (PHI). However, as this paper describes, de-identified data is exempt from this regulation based on the Privacy Rule, which regulates "limited data sets." A broader level of anonymization is required beyond the 3 categories of "protected health information," deidentified data," and "limited data sets." Data that is being used for "health care operations" by covered entities should be removed of all identifiers that are not expressly required. Deidentification standards need to be updated based on updated technology capabilities for reidentification
Department of Health and Human Services, Federal Coordinating Council for Comparative Effectiveness Research: (June 30, 2009)	Federal Coordinating Council for Comparative Effectiveness Research: Report to the President and Congress	Report	The report states that the primary investment for CER funding should be data infrastructure. Data infrastructure could include linking current data sources to enable answering CER questions, development of distributed electronic data networks and patient registries, and partnerships with the private sector. Secondary areas of investment are dissemination and translation of CER findings, priority populations, and priority types of interventions. The CER activity and investments should be coordinated across the Federal Government and avoid duplicative effort. In addition, the funding should complement and link to activities and funding in the private sector to maximize the benefits to the American people.

Source	Title	Type	Description of the grey literature product
Department of Health and Human Services, Office of the National Coordinator for Health Information Technology: (August 9, 2011)	Metadata Standards To Support Nationwide Electronic Health Information Exchange	Federal Register Notice	According to the Office of the National Coordinator for Health Information Technology at the Department of Health and Human Services (ONC), the use of metadata holds great promise and the adoption of metadata standards can help rapidly advance electronic health information exchange across a variety of different exchange architectures. As recommended by the HIT Standards Committee: ONC is considering the following standard set of patient identity metadata: Name, DOB, zip code, patient identifiers; ONC is considering the following standard set of provenance metadata: a tagged data element (TDE) identifier; a time stamp; and the actor, the actor's affiliation, and the actor's digital certificate; and ONC is considering the following standard set of privacy metadata as rwhich would include the following data elements expressed according to the requirements explained below—a "policy pointer" and content metadata elements, data type, and sensitivity.
Expressor Software: David Fenstermacher (November 2010)	Metadata: The Cornerstone of Tomorrow's Healthcare Information Management Systems.	White Paper	To allow health information technology to become more useful in the space of comparative effectiveness research, the paper states that metadata structures need to be implemented to provide meaning to the data and allow it to be interoperable with other systems. Data governance responsibilities should also be laid out to ensure that the data is managed appropriately in relation to contextual and physical metadata.
FasterCures (January 2011)	Still Thinking Research: Strategies to Advance the Use of Electronic Health Records to Bridge Patient Care and Research	Report	According to this report, the first step in the process of transitioning to using electronic records is the digitization of the records. Then, the data must be collected and stored in a standardized manner to maintain data quality. Data entry should be incentivized for physicians, and this should be supplemented by strengthening the relationship with investigators. Patient input should also be considered when entering data, related to health literacy and access to care. Allowing patients to understand how their PHI will be used for clinical research will make them more comfortable about sharing it, participating in research, and providing informed consent.
FasterCures: (May 2006)	Ensuring the Inclusion of Clinical Research in the Nationwide Health Information Network	Report	The article provides recommendations and an action agenda that resulted from a 2006 day-long, 50-participant workshop sponsored by FasterCures, AHRQ, and the National Center for Research Resources (NCRR) at the National Institutes of Health (NIH) that discussed how EHRs and clinical research can be made an integral part of the National Health Information Network (NHIN). Recommendations included creating data standardization measures once information has been digitized, promoting a culture of data privacy, and incentivizing clinicians to enter patient-centered data.
Federal Trade Commission: (February 2010)	Data Protection Accountability and the Appropriate Use of De-Identified Data	White Paper	This paper details an accountability-based approach to data governance that focuses on setting data protection goals based on legal requirements, public policy, self-regulation and best practices. Accountable organizations take responsibility for the data they safeguard by ensuring they have appropriate systems, policies and procedures, training, monitoring and oversight in place. Requiring patient consent for every data collection or use would cripple the flow of health data and impede quality research, healthcare improvement efforts and basic healthcare operations. A more rational method, and one advanced by our nation's health information privacy laws, is to undertake considerable efforts to ensure patient-identifiable information is not disclosed inappropriately and that data analysis and research is conducted with de-identified health information whenever possible.

Source	Title	Type	Description of the grey literature product
Georgetown University Health Policy Institute: Testimony by Joy Pritts (July 27, 2005)	Testimony before the United States House of Representatives Committee on Ways and Means Subcommittee on Health on Health Care Information Technology: Harmonizing Laws Governing the Confidentiality of Health Care Information	Congressional Testimony	The testimony focuses on what, if any, actions the federal government should take with respect to protecting the confidentiality of health information in order to facilitate the electronic exchange of health information, including the development of a national health information infrastructure (NHII). The speaker states that the HIPAA Privacy Rule is inadequate and that it is crucial that the privacy of health information not be compromised in the interest of expediency. Federal privacy protections for health information should be expanded to ensure that standards for using and disclosing health information are in place for everyone who receives or creates identifiable health information and that federal law should also ensure that those who improperly obtain use and disclose health information are subject to civil and criminal penalties.
Heritage Foundation: Helen Evans (February 4, 2009)	Comparative Effectiveness in Health Care Reform: Lessons from Abroad	White Paper	According to the paper, many industrialized countries have bodies charged with health technology assessments or comparative effectiveness studies. Despite this, the evolution of these bodies and their responsibilities at the national decision-making level has been far from uniform. It is only by returning health care to a genuinely patient-centered and consumer-driven health care marketplace that information, innovation, and best practice will permeate the complex array of health care arrangements in both the public and the private sectors. A comparative effectiveness strategy that relies on central planning and coercion would not only be counterproductive in the long run--because it would undermine the incentives for medical innovation--but would also lead to the imposition of cost constraints that would worsen patients' medical conditions and damage the quality of their lives.
Heritage Foundation: Tevi Troy (June 25, 2009)	Health Information Technology: The Case for a Sound Federal Policy	White Paper	According to the paper, there are still a number of unanswered questions about how to implement EHRs so that they create the maximum benefit for patients and the minimum disruption for America's already stressed health care system. There are a number of philosophical (political) and practical hurdles that must be overcome so that the implementation of this enterprise can be successful. The Administration should encourage the Department of Health and Human Services to: Aim Higher Than Electronic Cash Registers,Not Pick Winners and Losers, and Adopt a Platform Model Approach.
ID Experts: (November 2010)	2nd Annual Benchmark Study on Patient Privacy and Data Security	White Paper	This paper states that the number of data breaches among healthcare organizations participating in the 2010 and 2011 studies is still growing—eroding patient privacy and contributing to medical identity theft. Widespread use of mobile devices is putting patient data at risk. Despite policies and federal mandates, prevention of unauthorized access to patient information is not a priority in many organizations in this study. Diminished productivity and financial consequences for healthcare organizations can be severe when a data breach incident occurs. Medical identity theft poses a greater risk to patients.
Institute of Medicine: (June 2009)	Initial National Priorities for Comparative Effectiveness Research	Institute of Medicine Report	The committee recommends a balanced portfolio of research topics that, collectively, address broad societal needs. To evaluate a topic's importance, the committee formulated criteria that would identify not only those diseases and conditions with the greatest aggregate effect on the health of the U.S. population, but also less common conditions that severely affect individuals in vulnerable subgroups of the population. Among the high priority topics were interventions such as disease prevention, systems of care, drug therapies, devices, surgery, and monitoring of disease. The priority list includes 29 research areas , affecting a broad range of age and ethnicity. Twenty-four of the 100 topics affect special populations.

Source	Title	Type	Description of the grey literature product
Integrating Data for Analysis, Anonymization and SHaring (iDASH), University of California San Diego	Research and Development.	Website - Academic Center Summary	According to the website, iDASH enables wide-scale integration of diverse infrastructures, services, and tools for biomedical and behavioral investigators, regardless of their institutional affiliation, to widely engage in global collaborations. Moreover, iDASH allows scientists to concentrate on advancing their research while knowing that their underlying infrastructure — security, patient privacy, and data integrity — is not compromised. It provides an intuitive, innovative portal for accessing algorithms, open-source software, data storage, and training to facilitate secure, patient-anonymous data analysis and sharing.
Kroll Fraud Solutions: (April 2010)	2010 HIMSS Analytics Report: Security of Patient Data	White Paper	This paper describes the landscape of patient data security in the United States, and how this has changed between 2008 and 2010. The cost of a data breach is very high, and includes direct and indirect costs. Protecting patient data also improves patient satisfaction with the institution.
Lewin Group Center for Comparative Effectiveness Research: Clifford Goodman (October 28, 2009)	Comparative Effectiveness Research and Personalized Medicine: From Contradiction to Synergy	White Paper	This paper states that for CER to contribute to personalized medicine (PM), it must account for patient differences that influence the impact of interventions on health outcomes. Aligning CER and PM means that PM is subject to prevailing evidence requirements for screening, diagnostic, therapeutic, and other interventions. Full alignment of PM and CER depends on adoption of health information technology (HIT). CER is influencing innovation in PM, including enabling new opportunities and diminishing prospects for some less likely to fare well in a market informed by head-to-head comparisons. Communications and applications of CER findings and other evidence must be adaptive and targeted to clinicians, patients, payers, and the public.
Medicare Payment Advisory Commission (MedPAC): (June 2004)	Information technology in health care.	Report to Congress	Significant barriersto successfully implementing IT remain for many providers and the market forces encouraging adoption are weak. Current efforts may need to be expanded or new strategies developed to stimulate broader diffusion of health IT. Several legislative proposals, information technology experts, and research groups, such as the IOM, have suggested other ways to encourage faster adoption of IT. Options include: payment policy, loan funds, grants, and requirements to adopt specific technology. This review shows that IT use in health care is growing, but providers do experience barriers. Implementation is difficult, making the risk of investment high.
Mini-Sentinel Coordinating Center (sponsored by the U.S. Food and Drug Administration (FDA))	Background	Website	According to the website, the Mini-Sentinel pilot funds development of a single Coordinating Center that: 1) Provides the FDA a “laboratory” for developing and evaluating scientific methods that might be used in a fully-operational Sentinel System; 2) Affords the FDA the opportunity to assess safety issues using existing electronic healthcare data systems; and 3) Allows the FDA to learn more about the barriers and challenges to building a viable and accurate system of safety surveillance for FDA-regulated medical products. Key features include: active surveillance, collaboration, coordinating center, distributed data approach, data sources, rapid response to queries, communication of results, methods development, policy development, impact of FDA actions, and engagement with related efforts.
ModernMedicine: Jill Wechsler (April 2010)	FDA information systems, Sentinel Initiative seek more timely drug safety information	Article	The article describes the Sentinel Initiative, which was launched by the FDA in 2008, and now is establishing a “Mini-Sentinel” system that will tap into medical records held by large health plans and insurers. Legal experts want FDA to develop model procedures for when, how, and to whom to report drug safety findings as a way to set policies that can be applied in court. While Sentinel’s prime purpose is to serve FDA’s regulatory mission, the long-term vision is that it will be part of a larger health information system that also provides data for outcome studies, comparative effectiveness research, and health system quality reporting.

Source	Title	Type	Description of the grey literature product
National Cancer Institute: (November 2007)	caBIG Pilot Phase Report 2003-2007	Report	The report provides an overview of caGrid, a set of specifications and software modules that define a data transmission network upon which computer services operate to transmit data between collaborators. Key points include: caGrid was designed to include software features that ensure authorization and authentication of users and data security for any service operating on the grid; caBIG prepared for an initial response of 10-15 clinical centers, but had an actual response of 49 centers, which overwhelmed the system and affected funding; and there were some technology coordination gaps that needed to be addressed to achieve harmonization.
National Center for Policy Analysis: Devon M. Herrick, Linda Gorman, John C. Goodman (April 2010)	Health Information Technology: Benefits and Problems.	Policy Report	Although many proponents discuss the perceived benefits of HIT, missing from the debate is an honest discussion of experiences with actual HIT systems, and the obstacles and pitfalls of poorly designed systems. According to this report, the ultimate goal should be to improve quality, increase efficiency and add convenience — not just to create wired facilities. Policymakers should let the market, not the federal government, pick the technology that works best. Consumers should also have a say in the appropriate level of privacy that meets their needs.
National Institute for Health Care Reform: Emily Carrier, Hoangmai H. Pham and Eugene Rich (October 2010)	Comparative Effectiveness Research and Innovation: Policy Options to Foster Medical Advances	White Paper	According to the paper, the findings of CER studies likely will raise as many questions as they answer, and their results will be open to constant questioning and reinterpretation. However, well-designed CER policies can promote beneficial innovations and discourage development of treatments with relatively little benefit. The complexity and uncertainty of innovation may be such that no policy can ensure that every beneficial innovation is promoted and protected. This could prove a political weakness for CER unless thoughtfully addressed. Understanding consumers' beliefs about innovation and providing better access to participate in clinical trials of promising innovations might help address their concerns. CER policies that clearly state how access to innovation fits into societal values may help to amplify the effect of CER on patient and clinician decisions.
National Quality Forum (NQF): (December 2010)	Driving Quality—A Health IT Assessment Framework for Measurement.	A Consensus Report	Health IT use assessment can provide valuable information for most healthcare stakeholders, including the quality improvement community, the health IT vendor community, providers, payers, purchasers, and policymakers. The report states that the Health IT Utilization Assessment Framework provides a unique approach to identifying and measuring: 1) the use of health IT applications; 2) whether the workflow (driven by the system's user interface) occurs as designed; and 3) that such use improves care processes, quality, and safety. Standards development organizations, professional societies, workflow planners, and other key stakeholders and entities should collaborate to standardize, harmonize, and identify definitions of and gaps in roles for all users of clinical applications and health IT systems.
National Working Group on Evidence-Based Health Care: (August 2008)	The Role of the Patient/Consumer in Establishing a Dynamic Clinical Research Continuum: Models of Patient/Consumer Inclusion	Report	Evidence-based healthcare (EBH) is the concept of determining a patient's treatment by balancing scientific evidence, practitioner judgment, and patient experience and preference. This report states that the role of the patient/consumer in research extends beyond participation in the clinical trials. Next steps include defining a common language for evidence-based healthcare, including patients in all aspects of the research continuum, and evaluating the impact of patient engagement in research.

Source	Title	Type	Description of the grey literature product
New England Healthcare Institute: (September 2010)	From Evidence to Practice: Making CER Findings Work for Providers and Patients	White Paper	This paper states that the high priority now placed on dissemination of new CER findings reflects increased awareness that existing medical evidence of all types is haphazardly disseminated throughout the U.S. health care system, and that the uptake of new findings by clinicians and patients is protracted and uneven. Factors unique to the conduct of comparative effectiveness research create new hurdles for the dissemination and uptake of the research. CER dissemination policy should take full advantage of other trends in health care improvement, most particularly the deployment of health care information technology. Thoughtful and comprehensive dissemination policy will not only support the use of CER, but should go far to improve the utilization of all forms of scientific evidence in the health care system.
Observational Medical Outcomes Partnership: Points to Consider in Developing a Common Semantic Data Model and Terminology Dictionary for Observational Analyses: Patrick Ryan, Don Griffin, Luann Whittenburg, Dan Foltz, Marc Overhage (Last revised: March 3, 2009)	Points to Consider in Developing a Common Semantic Data Model and Terminology Dictionary for Observational Analysis.	White Paper	According to the paper, one of the goals of OMOP is portability of research methods across the Research Core data providers. The Common Data Model (CDM) facilitates this portability by managing the security and using ETL logic that creates data standardization. The CDM design used an Entity-Attribute-Value (EAV) modeling approach to data modeling, because this method places no artificial or arbitrary limits on the numbers or kinds of entities, attributes, or relationships that may be accommodated by the model. An Entity-Relational (ER) model approach is also being used for more static portions of the model.
RAND Corporation	A National Health Information Network – What Are the Real Privacy Issues?	RAND Health Fact Sheet	The fact sheet is based on Greenberg MD and Ridgely MS, "Patient Identifiers and the National Health Information Network: Debunking a False Front in the Privacy Wars," Journal of Health & Biomedical Law, Vol. 4, No. 1, 2008, pp. 31–68 and states that the controversy over Unique Patient Identifiers (UPIs) distracts from the key privacy issues connected with a National Health Information Network (NHIN): namely, the need to strengthen HIPAA privacy rules and to reconcile current state laws on health information privacy. Many reforms to increase health privacy have been suggested including the following: extend HIPAA privacy rules; enact federal legislation against misuse of personal health information; enact federal rules to govern operation of a NHIN as well as strong enforcement procedures; build privacy protection into a NHIN architecture.
RAND Corporation	Analysis of Comparative Effectiveness	RAND Technical Report	This report states that patient experience would improve if comparative effectiveness research were incorporated into shared decision-making. There is no clear evidence about how such research would affect spending, consumer financial risk, waste, or reliability of care. Theory suggests that comparative effectiveness research could improve health if it drives payers, providers, and patients toward more beneficial treatment options. Such research is not relevant to coverage. Establishing a national center for comparative effectiveness research would be easy; translating research into better clinical decisionmaking and increased health system efficiency would be complex.
Recombinant by Deloitte	i2b2	Website	The website describes i2b2 (Informatics for Integrating Biology & the Bedside), an open-source platform for de-identified cohort discovery, and for managing and delivering clinical data sets for research with appropriate IRB approval. An i2b2 implementation consists of a data mart of clinical, research, and administrative data, and an interface to construct and manage queries and data sets. Researchers can use i2b2 to perform self-service queries of de-identified data for HIPAA-compliant cohort discovery and hypothesis testing. Plug-ins and extensions to i2b2 are available to add additional data capture, extraction, and analysis functionalities to the platform's core capabilities.

Source	Title	Type	Description of the grey literature product
Robert Wood Johnson Foundation and Urban Institute: Elizabeth Docteur and Robert Berenson (February 2010)	How Will Comparative Effectiveness Research Affect the Quality of Health Care?	White Paper	According to the paper, there are significant challenges in undertaking a comparative effectiveness (CE) initiative and understandable concerns about CE's having unanticipated and undesirable impacts. Efforts to distill lessons from extensive past experience in federal work on comparative effectiveness, now extending over at least three decades, should be a priority. While investing in CE can be a path for improving the quality of health care and increasing the value of health expenditure, just doing the research is not enough to change practice. Rather, CE should be considered a valuable part of a larger effort to foster evidence-based medicine, along with changes in incentives and the organization of health-care delivery that are essential to promote and support high-quality health care.
Robert Wood Johnson Foundation : (January 18, 2007)	A 'Rapid-Learning' Health Care System.	White Paper	A rapid-learning healthcare system uses health information technology to collect data that evaluates the effectiveness of various medical interventions in different situations. According to the paper, this system should be developed with the goal of interoperability among different servers.
Robert Wood Johnson Foundation : (June 2010)	How Registries Can Help Performance Measurement Improve Care	White Paper	The paper outlines how to improve care and advance performance measurement with registries data. Short-term recommendations include implementing nationally-endorsed measures based on claims data, deploying home-grown registry functionalities in national registries, clarifying applications of HIPAA to data linkage activities, and leveraging demographic data in registries to assess disparities. Long-term recommendations include data element and definition standardization, creation of patient identity management methods, allowing interoperability between registries and EHRs, standardization of data quality assurance and risk adjustment methods, standardization of linkage methods, encouragement of provider participation in registries, and the creation of sustainable business models for registry programs.
Robert Wood Johnson Foundation: (July 1, 2011)	Does Health Information Technology Improve Quality of Care?	Health Policy Issue Brief	The issue brief states that while many industries have lowered costs and improved quality through investments in information technology, the health care sector has been slow to follow suit. Health information technology can improve communication: HIT can ease communication between patients and doctors, which can result in better outcomes. Many EHR systems include online patient portals in which patients can view test results, see aspects of their medical records and email their doctors. Health information technology can drive efficiency gains: Although adopting a new EHR system can be costly and time-consuming for staff, the long-term gains can be many. Information technology can improve the quality of care patients receive by averting medical errors, improving communication and boosting efficiency.
Robert Wood Johnson Foundation: Deven Mcgraw (January 1, 2009)	Legal Solutions in Health Reform: Privacy and Health Information Technology	White Paper	According to this paper, the perceived "gaps" in current federal legal protections for health information can be grouped into four categories: 1) who is covered; 2) what is covered; 3) state law variation; and 4) insufficient comprehension of and compliance with privacy protections. The solutions range from amending existing law or regulation to encouraging private action through market or other incentives. While there is consensus that efforts to facilitate widespread adoption and use of health information technology must move forward with appropriate protections for privacy and security, achieving consensus on the details of what privacy and security measures need to be put in place continues to be a challenge.
Sterling Health IT	A Universal Exchange Language Supports Comparative Effectiveness & Biomedical Research	White Paper	Universal Exchange Language (UEL) is a standardized framework that allows translation of a number of different sources into one vocabulary. According to the paper, a UEL could solve the problem of the lack of a national patient identifier, correct interoperability and portability issues, and combine free text and coded portions of the EMR.

Source	Title	Type	Description of the grey literature product
U.S. Chamber of Commerce: (January 2009)	The Health Care Comparative Effectiveness Toolkit: Promoting Value for Employee Health	White Paper	The report states that improving employee health management is one of the dominant topics for employers for the next decade. Comparative effectiveness information is only helpful if it is integrated into broader strategies of patient engagement, health care literacy, and employee health care support. The Chamber of Commerce recommends: (1) where relevant, engage patients and provide the Agency for Healthcare Research and Quality (AHRQ) effectiveness summaries, (2) encourage providers associated with your health plans to fully consider publicly available, evidence-based practice guidelines which are described below, and (3) begin to consider other ways to incorporate this growing base of information to improve coverage and employee health management
White House, PCAST: Executive Office of the President President's Council of Advisors on Science and Technology (December 2010)	Realizing the Full Potential of Health Information Technology to Improve Healthcare for Americans: The Path "Forward.	Report	The report discusses how information technology can be transformative in the health care field, including improving healthcare outcomes and reducing cost. It can also facilitate the development of personalized medicine, and allow patients to become more involved in their own health care. The considerations that have been proposed here in order to facilitate implementation include creating a universal exchange language for data standardization and sharing, ensuring that the EHRs fit into the clinical workflow, and promoting the surveillance and research value in utilizing the system.
Wiki - eNotes	caBIG	Website	According to this website, caBIG sought to provide foundational technology that enables a new approach to biomedicine called a "learning healthcare system." This model of research and care delivery relies on the rapid exchange of information between all sectors of research and care, so that researchers and clinicians are able to collaboratively review and accurately incorporate the latest findings into their work. The ultimate goal is to speed the biomedical research process, leading to improved patient outcomes and more efficient healthcare delivery. Although the goal was considered laudible, much of the software was unevenly adopted after being developed at great expense to compete with commercial offerings.
Wiki - Health Informatics	Term: Strategic Health IT Advanced Research Projects (SHARP) Program	Website	This website provides detailed information about the SHARP program, a research program that has been awarded $60 million and is funded by The Office of the National Coordinator for Health Information Technology (ONC). Awardees in the SHARP Program implement a focused research project in one of the following four areas where breakthrough advances are needed to address barriers to the adoption of health IT to meet the goal of making electronic health records (EHRs) available for all Americans by 2014. 1. Security of Health Information Technology 2. Patient-Centered Cognitive Support 3. Healthcare Application and Network Platform Architectures 4. Secondary Use of EHR Data.
Wiki - Internet2	Guidelines for Data De-Identification or Anonymization	Website	According to the website, before embarking on a data de-identification project, high-level challenges and risks must be identified to determine how to appropriately mitigate risks in the context of the proposed use of the data. These challenges and risks include: No regulation of de-identified data, lack of clear definition of de-identified or anonymous data, paper-based vs. electronic data, types of de-identified or anonymous data, need for re-identification and careful use of re-identification keys, balancing risk with value, handling and use considerations, data classification, international considerations and providing services for de-identifying or anonymizing data.

Appendix C: EDM Forum – Grey Literature Abstraction Form

Citation: Vancouver

Level of Review

Reader: Date:

Tagged Category: *Select one theme that best represents this article

- “Cloud Computing”
- “Cohort Identification”
- “Data Use and Quality”
- “De-Identification”
- “Governance”
- “Institutional Review Board”
- “Learning Healthcare system”
- “Library of Phenotypes”
- “Metadata”
- “Natural Language Processing”
- “Patient Involvement”
- “Overview”
- “Research Networks”
- “Security”
- “Single Point Access”
- “Standardized Data Collection”

Extracted Article Information

Primary Objective/Purpose/Argument/Aims

Resource Type *Select one

- Web Page
- Working Paper
- Newsletters
- Surveys
- White Paper
- Thesis/Dissertation
- Committee Reports and Memoranda
- Other

Resource Notes (describe specifically what the resource topic and content are, e.g. policy paper, “About” page, methodology section, etc.)

Methods (if applicable, note N/A if not)

Major Themes/Issues

Key quotes/Data Sources

Vocabulary/Key Concepts/Definitions

Results/Conclusions (if applicable)

Author’s Perspective (if applicable)

Funding Sources/RFA # (if applicable)

Analysis of the Literature

EDM Forum Relevance Rating (1 = no relevance, 5 = extremely relevant), please explain the strengths and weaknesses of the article.

Quality Rating (1 = lowest quality, 5 = highest quality), please explain the strengths and weaknesses of the article (e.g. level of review, citation, references).

List any ideas for topics of discussion, questions, or activities for the EDM Forum you gained from reading this article.

From the references (if applicable), list any articles to add to the Grand Round readings or authors to engage in EDM Forum activities.

Flag as potential author/contributor to EDM Forum deliverables? *Y/N checkbox with optional free-form response

EDM Forum Relevance Rating (1 = no relevance, 5 = extremely relevant), please explain the strengths and weaknesses of the article.

Endnotes

1. Hamilton Lopez M, Holve E, Sarkar IN, Segal C. Building the Informatics Infrastructure for Comparative Effectiveness Research (CER): A Review of the Literature. Med Care. 2012 Jul;50 Suppl:S38-48.

2. The Fourth International Conference on Grey Literature (GL '99) in Washington, DC, in October 1999 defined grey literature as follows: "That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers." Alberani defines grey literature as publications that include, but are not limited to the following types of materials: reports (pre-prints, preliminary progress and advanced reports, technical reports, statistical reports, memoranda, state-of-the art reports, market research reports, etc.), theses, conference proceedings, technical specifications and standards, non-commercial translations, bibliographies, technical and commercial documentation, and official documents not published commercially (primarily government reports and documents). EDM Forum staff drew on both definitions to identify grey literature on ECD.

3. Blackhall K, Ker K. Finding studies for inclusion in systematic reviews of interventions for injury prevention – the importance of grey and unpublished literature. Inj Prev. 2007 October; 13(5): 359.

4. AcademyHealth (Producer). (2011). 101 - 103: Grey Lit [Video webcast]. In: Grey Literature Web Conference Series. Retrieved from http://www.academyhealth.org/Training/ResourceDetail.cfm?itemnumber=6670

5. Committee on Standards for Systematic Reviews of Comparative Effectiveness Research, Institute of Medicine. "Front Matter." Finding What Works in Health Care: Standards for Systematic Reviews. Washington, DC: The National Academies Press, 2011.

6. When searching two terms together, we employed the Boolean search technique (AND), enabling us to narrow the search.

7. Singer Cohen R. Master List of Grey Literature-Producing Organizations. AcademyHealth. 2011. Internally-generated and used document, available from project staff upon request. Funded by the National Library of Medicine.

8 Consultants who advised on the grey literature process: Rebecca Singer Cohen organized the grey literature webinar series for AcademyHealth, and was involved in the creation of the master list of grey literature-producing organizations. The search strategy was discussed in detail, including the two lists and how they were developed. She suggested these resources, as well as experts that could be consulted for further review, including Ione Austan.

   Ione Austan, M.L.S., of the National Information Center on Health Services Research and Health Care Technology (NICHSR), a component of the National Library of Medicine, has a background in Library Science. After creating an initial grey literature search strategy and list of organizations for review for this purpose, Ione was asked to review these materials and provide comments on the search methods, as well as the list of organizations that were chosen. She provided several additional sources for inclusion in the review.

   Neil Sarkar is a clinical informaticist and the library consultant for the EDM Forum. He reviewed the document and provided additional resources for review.

9 Hepworth, N., Hooper, V., Hellebrandt, D., Zeitoun, M., Lankford, B., & Pegram, G. What factors determine the performance of institutional mechanisms for water resources management in developing countries in terms of delivering pro-poor outcomes, and supporting sustainable economic growth? 2011 CEE protocol 11-006. Collaboration for Environmental Evidence: www.environmentalevidence.org/SR11006.html.

10 Bielska IA. Using population health surveys to measure the use of services and prevalence of psychiatric and/or behavioural conditionsl in individuals with intellectual disability. 2009. A thesis submitted to the Department of Community Health & Epidemiology in conformity with the requirements for the degree of Master of Science Queen's University. Kingston, Ontario, Canada

11 In the cases that the documents/webpages provided historical context or frameworks for using ECD for research, they were coded as "overview."

12 To the extent that the reviewers did not identify any relevant articles that align with the concepts of interest, they assessed these as being 'gaps' in the grey literature. In some instances, each of these concepts were discussed as a challenge or issue in the context of papers on other topics, but were not the primary focus of the article. In the grey literature review, no papers were identified as focusing on "natural language processing", "IRB", "library of phenotypes", "governance", "single point access" to research data, "cloud computing" or "cohort identification."

13 The transience of available material was discussed during a Health Services Research Information Advisory Committee Meeting held on April 27, 2012. In particular, the group acknowledged the challenges inherent in a body of work that is less traditional "literature" and instead freed information produced in many different formats, by many different organizations, for many different audiences.