

## ORIGINAL RESEARCH



# Actualizing Personalized Healthcare for Women Through Connected Data Systems: Breast Cancer Screening and Diagnosis

通过连接的数据系统推行适用于女性的个性化医疗护理：乳腺癌筛选和诊断

Puesta al día de la atención médica personalizada para las mujeres mediante sistemas de datos conectados: detección y diagnóstico del cáncer de mama

Heidi D. Nelson, MD, MPH, *United States*; Roshanthi Weerasinghe, MPH, *United States*

## Author Affiliations

Providence Cancer Center, Providence Health & Services Oregon (Dr Nelson, Ms Weerasinghe); Departments of Medical Informatics and Clinical Epidemiology and Medicine, Oregon Health & Science University (Dr Nelson), Portland.

## Correspondence

Heidi D. Nelson, MD, MPH  
heidi.nelson@providence.org

## Citation

Global Adv Health Med. 2013;2(5):30-36. DOI: 10.7453/gahmj.2013.054

## Key Words

Women's health, breast cancer, screening, diagnosis, data systems, personalized healthcare

## Disclosures

The authors completed the ICMJE Form for Disclosure of Potential Conflicts of Interest and disclosed none.

## ABSTRACT

**Background:** Healthcare organizations have invested in electronic patient data systems, yet use of health data to optimize personalized care has been limited.

**Primary Study Objective:** To develop and pilot an integrated source of health system data related to breast healthcare.

**Methods/Design:** This study is a quality improvement project. Patient-level data from multiple internal sources were identified, mapped to a common data model, linked, and validated to create a breast healthcare-specific data mart. Linkages were based on matching algorithms using patient identifiers to group data from the same patient. Data definitions, a data dictionary, and indicators for quality and benchmarking aligned with standardized measures. Clinical pathways were developed to outline the patient populations, data elements, decision points, and outcomes for specific conditions.

**Setting:** Electronic data sources in a community-based health system in the United States.

**Participants:** Women receiving breast cancer screening, prevention, and diagnosis services.

**Main Outcome Measures:** Distribution of mammography examinations and pathologic results of breast biopsies.

**Results:** From 2008 to 2011, 200 768 screening and 50 200 diagnostic mammograms were obtained; rates varied by age over time. Breast biopsies for 7332 women indicated 23.3% with

invasive breast cancer, 6.7% with ductal carcinoma in situ, and 70.0% with nonmalignant diagnoses that would not have been further differentiated by administrative codes alone.

**Limitations:** Evaluation of validity and efficiency and additional tracking of clinical outcomes are needed.

**Conclusions:** The creation of a patient-centered data system by connecting and integrating disparate data sources within a large health system allows customized analyses of data and improves capacity for clinical decision making and personalized healthcare.

## 摘要

**背景:** 医疗护理组织已经在电子患者数据系统方面进行了大量投入, 但将健康数据用于完善个性化护理却一直受限。

**主要研究目标:** 建立一个与乳房医疗护理相关的健康系统数据综合性来源, 并进行试点试验。

**方法/设计:** 本研究是一个品质提升项目。我们对来自于多个内部来源的患者数据进行识别、将其映射到一个普通的数据模型上、进行连接和验证, 进而创建一个乳房医疗护理专用的数据集市。连接是基于匹配运算法则进行的, 即将采用患者识别码的运算法则与来自同一患者的分类数据进行匹配。使数据定义, 数据字典以及品质和基准测试指标与标准化衡量指标保持一致。最后, 制定临床路径, 概述具体条件下的患者群体、数据元素、决策点和结果。

**环境:** 电子数据源位于美国境内一个基于社区的健康系统中。

**参与者:** 正在接受乳腺癌筛选、预

防和诊断服务的女性。

**主要结果测量指标:** 乳房 X 线照相术检查的分布情况和乳腺活组织检查病理结果。

**结果:** 从 2008 年至 2011 年, 已经拍摄了 200,768 张筛选和 50,200 张诊断乳房 X 线照片; 速率随年龄和时间的不同而发生变化。针对 7332 名女性进行的乳腺活组织检查表明, 23.3% 罹患浸润性乳腺癌, 6.7% 罹患导管原位癌, 70.0% 为非恶性诊断, 并且无法单独通过管理代码进行进一步区分。

**限制:** 需要对有效性和疗效进行评价, 并对临床结果进行进一步跟踪。

**结论:** 将不同的数据源连接并整合至一个大型的健康系统中, 进而创建一个以患者为中心的数据系统, 这有助于进行个性化的数据分析并增强临床决策和个性化医疗护理的能力。

## SINOPSIS

**Antecedentes:** Las organizaciones sanitarias han invertido en sistemas electrónicos de datos de pacientes, pero el uso de los datos de salud para optimizar la atención personalizada ha sido limitado.

**Objetivo principal del estudio:** Desarrollar y poner a prueba una fuente integrada de datos del sistema de salud relacionados con el cuidado médico del pecho.

**Métodos/Diseño:** Este estudio es un proyecto de mejora de la calidad. Se identificaron los datos a nivel de paciente de múltiples fuentes internas, se asignaron a un modelo común de datos, se conectaron y se validaron para crear un mercado de datos específico para el cuidado

médico del pecho. Las conexiones se basaron en algoritmos de correspondencia que utilizaban identificadores de pacientes para agrupar los datos de la misma paciente. Se alinearon definiciones de datos, un diccionario de datos e indicadores de calidad y de referencia con mediciones normalizadas. Se desarrollaron vías clínicas para definir las poblaciones de pacientes, los elementos de los datos, los puntos de decisión y los valores para afecciones específicas.

**Entorno:** Fuentes de datos electrónicos en un sistema de salud comunitario en los Estados Unidos.

**Participantes:** Mujeres que recibían servicios de detección, prevención y diagnóstico del cáncer de mama.

**Criterios de valoración principales:** Distribución de exploraciones mamográficas y resultados patológicos de biopsias de mama.

**Resultados:** Entre 2008 y 2011 se obtuvieron 200 768 mamografías exploratorias y 50 200 mamografías diagnósticas; las tasas eran diferentes según la edad en el curso del tiempo. Las biopsias de mama de 7332 mujeres indicaron un 23,3 % con cáncer de mama invasivo, un 6,7 % con carcinoma ductal in situ y un 70,0 % de diagnósticos benignos,

que no se habrían diferenciado con más detalle solo por medio de códigos administrativos.

**Limitaciones:** Son necesarios una evaluación de la validez y la eficacia y un seguimiento adicional de los resultados clínicos.

**Conclusiones:** La creación de un sistema de datos centrado en el paciente por medio de la conexión e integración de fuentes dispares de datos existentes en un amplio sistema de salud permite análisis personalizados de los datos y mejora la capacidad para la toma de decisiones clínicas y para la atención médica personalizada.

## BACKGROUND

Many healthcare organizations and institutions have invested heavily in electronic patient data systems, yet use of health data to optimize personalized prevention, diagnostic, and therapeutic care has been limited. Although many factors contribute to this missed opportunity, the quality, completeness, and accessibility of health system data are major concerns.<sup>1,2</sup> Often, health systems purchase data software from multiple vendors for the primary purposes of billing and scheduling, rather than for patient care and evaluation. As a result, multiple nonlinked systems may coexist within a health system but fail to connect essential patient-centered information. Health systems tangled in this data web may require extensive reengineering to effectively access patient data for personalized care, while those early in their development of electronic data systems have an opportunity to build on the successes and avoid the shortcomings of existing systems.

Much of the published evidence of the effectiveness of health information systems to improve healthcare has focused on specific components of care. These include medication management (orders, reminders, adverse events, alerts); preventive care reminders and adherence to guidelines; diagnostic aids; chronic disease management; health outcomes; efficiency and cost; and satisfaction.<sup>1,3</sup> Studies of these systems in western countries indicate improved care with drug ordering and preventive care reminders but not for health outcomes, resource utilization, or cost, although the lack of effect relates to shortcomings of the studies as well as to the systems themselves.<sup>1,3</sup> Characteristics of effective information systems include using in-house systems, developers as users, integrated decision support, and benchmark practices.<sup>1</sup> In addition, effective systems address contextual issues related to patients and providers, incentives, interoperability, implementation, improvement, and policies.<sup>1</sup>

Despite the lack of evidence to guide the development of effective health information systems, the need

to access and fully utilize patient data remains. Connecting disparate data sources to individual patients and across time is an initial step in improving personalized healthcare services and health outcomes, particularly for conditions associated with fragmented care. Breast cancer screening, prevention, and diagnosis provide an example that is important to women. These services often are disjointed and subject to practice variation in the United States.<sup>4,5</sup> In addition, essential patient data on mammography and breast procedures often are difficult to extract from radiology and pathology data sources. For example, details of diagnostic breast procedures and pathology diagnoses are generally available only by manually reviewing reports. Diagnosis Related Group (DRG) Current Procedure Terminology or other coding systems indicate only if a procedure resulted in a diagnosis of invasive cancer, ductal carcinoma in situ (DCIS), or nonmalignant lesion. Additional diagnoses are usually embedded within the text fields of dictated pathology reports, limiting their access.

Though breast cancer is considered common in the United States, most women will never have it.<sup>6,7</sup> Early detection through screening continues to be an important effort to reduce breast cancer morbidity and mortality, and all women are eligible for breast cancer screening and prevention services over several decades of their lives.<sup>4,5</sup> Although barriers to services exist for many women, screening ideally begins with periodic mammography at age 40 or 50 and continues every year or two for 25 years or more.<sup>4,5</sup> Approximately 9 to 12 per 1000 women require breast biopsies because of suspicious radiographic lesions after one course of mammography, depending on the woman's age.<sup>8</sup> Women with physical findings such as breast lumps or skin changes also require breast biopsies. Many will require multiple biopsies during their lifetimes, although accurate cumulative estimates are not available.

Fortunately, for most women, biopsy results do not usually indicate breast cancer. However, simply having had a benign breast biopsy is associated with increased

breast cancer risk.<sup>9</sup> Several pathological types are considered high-risk lesions, including carcinoma in situ, atypical hyperplasia, and other atypical types. For example, lobular carcinoma in situ (LCIS), atypical ductal hyperplasia (ADH), and atypical lobular hyperplasia (ALH) increase 10-year breast cancer risk to 17% to 26%.<sup>10</sup> The increased detection of high-risk lesions in recent years is related to higher rates of mammography screening and subsequent biopsies of suspicious findings. While screening provides opportunities to identify and reduce cancer risks, it also increases healthcare burdens for women, health systems, and payers.

Although thousands of women are diagnosed with nonmalignant high-risk breast lesions each year, choosing the optimal prevention and diagnostic options is difficult and practice varies. Studies of the effectiveness of clinical management options are lacking. Women at increased risk of breast cancer may require earlier and more frequent mammography, as well as additional imaging modalities, such as breast ultrasound and magnetic resonance imaging (MRI), compared to women at average risk.<sup>5,11</sup> In addition, recommendations advise referral of women with significant family histories of breast cancer for genetic counseling and if eligible, to receive genetic testing<sup>12</sup> and be considered for risk reduction medications<sup>13,14</sup> or surgery. However, women often are unsure about what screening and prevention services would be appropriate for them because services are fragmented, lack uniform standards,<sup>15</sup> and are not patient centered.

The purpose of this project is to develop and pilot an integrated and more clinically useful source of health system data related to breast healthcare within a large, community-based health system. The ultimate goal is to use these data to optimize delivery of personalized health care. This study serves as a case example, presenting an approach that would be relevant to other conditions and settings. It focuses on breast cancer screening, prevention, and diagnostic services because of its volume, complexity, fragmentation, and opportunity for improvement and because it connects patient data from existing data sources that currently are difficult to access.

## METHODS

### Study Design

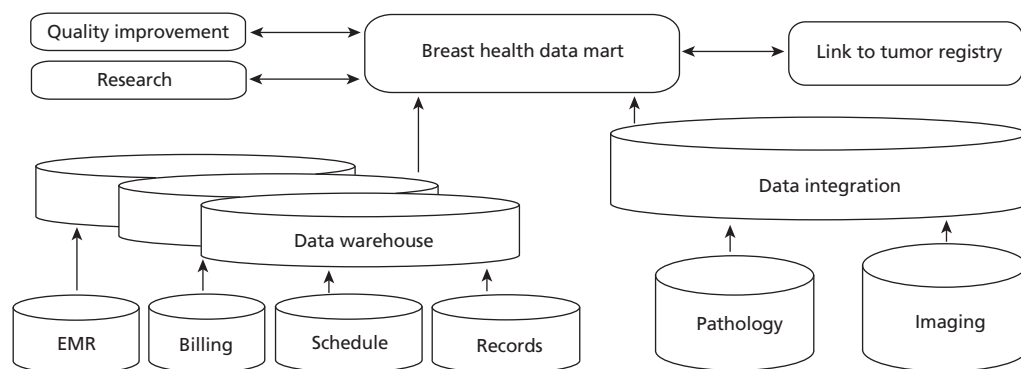
This study was designed as a quality-improvement project at a large, nonprofit community health system in the United States. A health system clinical team was assembled to guide the project and assure its clinical relevance while working with the health system's in-house technology team. The clinical team included health system experts in radiology, pathology, surgery, oncology, primary care, tumor registry, and informatics. Meeting periodically, the team worked collaboratively to create work plans, assess progress, provide clinical updates, and determine data priorities. The project was approved by the Providence Health and Services Institutional Review Board and Privacy Board.

### Participants

This project included patients receiving services at Providence Health & Services Oregon (Portland), an integrated health system of eight community hospitals and affiliated outpatient facilities across the state. The health system provides comprehensive care for breast cancer and related conditions, including screening, prevention, diagnosis, treatment, and survivorship care. Patients closely match the demographic and socioeconomic profiles of their communities, including women from inner city, urban, suburban, and rural areas, and many uninsured patients. The volume and diversity of Providence patients and the uniqueness of the health system's open-access policies provide an exceptional opportunity for community-based research.

### Procedures

Patient-level encounter data from multiple internal sources were identified, mapped to a common data model, linked, and validated to create a breast health-care-specific data mart. Key information sources included demographic and procedural data from administrative data sources and electronic medical records, imaging data from the radiology data system, and pathology data from the laboratory information system (Figure 1). Data from primary sources were integrated and stored in the clinical data warehouse,



**Figure 1** Health system primary data sources.

Abbreviation: EMR, electronic medical records.

and a subset of data was extracted to create the breast health-specific data mart. The data mart contains disease-specific data tables, providing an agile data source for customized ad hoc queries. The database structure was designed to interface with current and new data sources, and it undergoes continuous evaluation of the data extract, transform, and load process; data mapping; and validation as health system data sources are changed or upgraded.

Data from the various primary sources were linked based on matching algorithms using a number of patient identifiers to group data from the same patient. The health system uses a Master Patient Index (MPI) for each unique patient who accesses its extensive clinical network of clinics or hospitals. The data from these encounters are collected in the data warehouse. The MPI and other unique patient identifiers can be used to re-identify patients who leave and re-enter the system.

Currently, the data mart involves collection and analysis of existing documents, records, and pathological or diagnostic specimen data that are obtained as part of routine patient care. Additional existing data sources, such as the tumor registry, and potential new sources of quality improvement data, research data, or data provided directly by patients can be linked to the data mart in the future. Patient privacy is maintained by avoiding direct identifiers, such as patient names and social security numbers. Security measures include highly secure and limited access to the database, encryption technology, and use of de-identified data sets for analyses.

The development of data standards for the data mart was based on an adaptation of the common data model (Figure 2). This model can be used to minimize variability and enable common interpretation from multiple data sources. It has been developed and used for drug safety research using large observational data sources, including work by the Observational Medical Outcomes Partnership (OMOP).<sup>16</sup> The data standards themselves were modeled after the Breast Cancer Surveillance Consortium (BCSC),<sup>17</sup> a National Cancer Institute research collaborative, and the National

Accreditation Program for Breast Centers (NAPBC), which has published National Quality Measures for Breast Centers.<sup>18</sup> Pathology data were categorized from an electronic pathology database of dictated reports using a standardized lexicon (The Breast Pathology Assessment Tool and Hierarchy for Diagnosis [BPATH-Dx]).<sup>19</sup> Data definitions, a data dictionary, and indicators for quality and benchmarking align with these various measures.

Clinical pathways were developed by the clinical team to outline the patient populations, data elements, decision points, and outcomes for specific conditions. Pathways can be useful in designing a data mart and for subsequent analyses. For a woman with a nonmalignant breast lesion, the clinical pathway generally involves a series of sequential steps including (1) detection of a suspicious finding through routine or diagnostic mammography or physical examination; (2) biopsy and identification of the finding; (3) prevention, diagnostic, and therapeutic decisions that vary according to the pathological lesion, practice patterns, and preferences; and (4) monitoring for subsequent outcomes (Figure 3). For example, most clinicians would recommend surgical excision of ADH and LCIS and 6-month follow-up mammography for any biopsy.<sup>15</sup> However, many steps in the pathway are subject to practice variation, and data have not been available to evaluate current practices.

### Measures and Analysis

Data describing the distribution of mammography examinations in the health system (by age, year, and indication) and pathological results of breast biopsies (by age, year, and type) served as the main outcome measures for piloting the integrated data mart. These measures were selected because of their clinical importance, the existence of standardized data definitions, and their requirements for data linkages from various data sources. Analysis included descriptive statistics including proportions.

### RESULTS

Using these approaches, initial analysis of the data mart indicated that during a 4-year period (2008 to

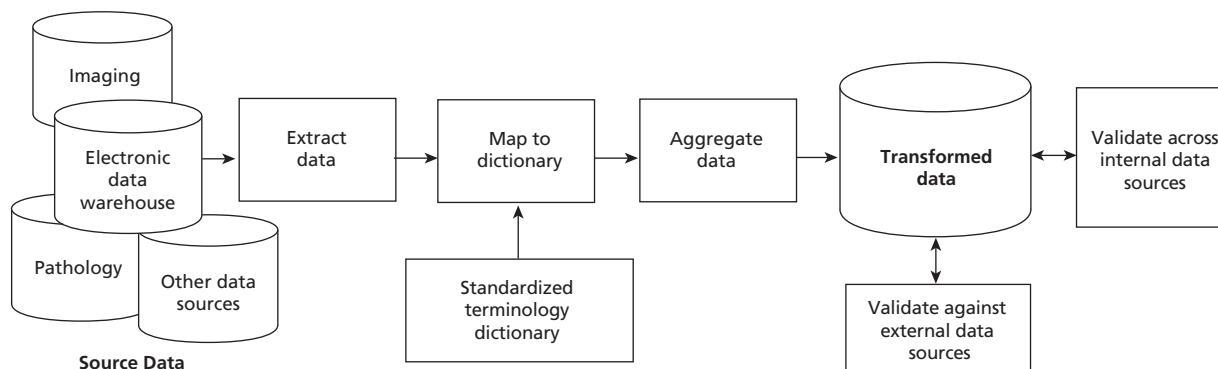
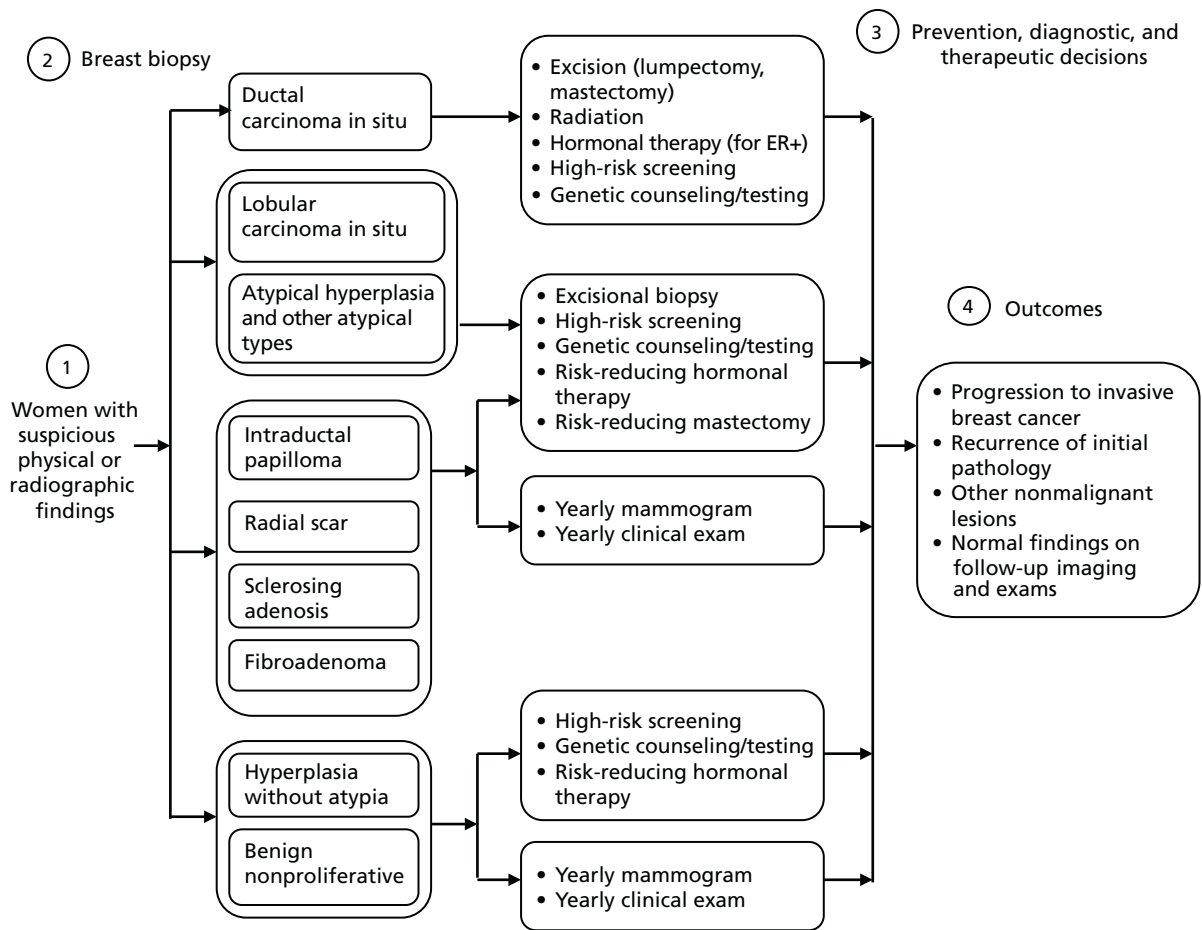


Figure 2 Common data model.



**Figure 3** Clinical pathway for nonmalignant breast lesions.

Abbreviation: ER+, estrogen receptor positive.

2011), 200768 screening and 50200 diagnostic mammograms were obtained in the health system (Table 1). Mammography rates varied by age and other characteristics over time, such as insurance status.<sup>20</sup>

During this period, breast biopsies for 7332 women were evaluated by health system pathologists. Using the diagnostic hierarchy to query the electronic pathology database, the data mart categorized patients into discrete diagnostic groups based on the most clinically significant diagnosis from each specimen, consistent with clinical practice. Results indicated that 23.3% (1709) of patients had invasive breast cancer, 6.7% (491) had DCIS, and 70.0% (5132) had nonmalignant diagnoses that would not have been further differentiated by administrative DRG codes alone (Table 2). These included 336 cases of ADH, 89 cases of ALH, 82 cases of LCIS, and 106 cases of other types of atypia (flat epithelial atypia, papillary atypia, apocrine atypia).

**CONCLUSIONS**

The creation of a patient-centered data system by connecting and integrating disparate data sources within a large health system allows customized analyses of data. For breast care services, patient-specific data were linked across patient records and administrative, radi-

ology, and pathology data sources. This type of data integration improves the health system’s capacity to provide individual-level data to support clinical decisions and actualize personalized healthcare.

Next steps require connecting the women’s biopsy results with data about their prevention, diagnostic, and therapeutic decisions in the clinical pathways and then linking them to long-term outcomes. In addition, associations between patient characteristics and outcomes can be determined in order to identify potential predictors for specific outcomes. These characteristics include age, social and demographic variables, family history of breast cancer, other known risk factors for breast cancer, results of previous imaging studies and biopsies, and clinical breast findings (eg, symptoms, palpable mass and size, etc), among others.

The data mart also requires further evaluation of its validity and efficiency. While validity is generally defined as a measure of the degree of erroneous and missing data, there are no established standards to evaluate the efficiency of clinical data sources. Efficiency can be determined by its accessibility (ie, how easy it is to run queries and pull specific data elements from data stores), format translation (ie, how much translation is required to standardize the data elements in the common data

**Table 1** Mammography Examinations in the Health System 2008-2011, n (%)

| Screening    |              |              |              |              |               |
|--------------|--------------|--------------|--------------|--------------|---------------|
| Age, y       | 2008         | 2009         | 2010         | 2011         | Total         |
| <40          | 1226 (2.4)   | 995 (2.0)    | 686 (1.4)    | 661 (1.3)    | 3568 (1.8)    |
| 40-49        | 11934 (23.7) | 11179 (22.3) | 10173 (20.9) | 10965 (21.2) | 44251 (22.0)  |
| 50-59        | 15316 (30.5) | 15082 (30.1) | 14979 (30.7) | 15665 (30.3) | 61042 (30.4)  |
| 60-69        | 11529 (22.9) | 12292 (24.6) | 12800 (26.3) | 14081 (27.2) | 50702 (25.3)  |
| 70-79        | 6709 (13.3)  | 6864 (13.7)  | 6769 (13.9)  | 7118 (13.8)  | 27460 (13.7)  |
| ≥80          | 3584 (7.1)   | 3651 (7.3)   | 3306 (6.8)   | 3204 (6.2)   | 13745 (6.8)   |
| <b>Total</b> | <b>50298</b> | <b>50063</b> | <b>48713</b> | <b>51694</b> | <b>200768</b> |
| Diagnostic   |              |              |              |              |               |
| Age, y       | 2008         | 2009         | 2010         | 2011         | Total         |
| <40          | 1303 (11.2)  | 1459 (11.2)  | 1209 (9.7)   | 1229 (9.4)   | 5200 (10.4)   |
| 40-49        | 3334 (28.6)  | 3640 (28.1)  | 3491 (28.1)  | 3610 (27.5)  | 14075 (28.0)  |
| 50-59        | 3197 (27.4)  | 3481 (26.8)  | 3455 (27.8)  | 3626 (27.6)  | 13759 (27.4)  |
| 60-69        | 2151 (18.4)  | 2458 (18.9)  | 2475 (19.9)  | 2811 (21.4)  | 9895 (19.7)   |
| 70-79        | 1094 (9.4)   | 1239 (9.5)   | 1187 (9.5)   | 125 (9.5)    | 4772 (9.5)    |
| ≥80          | 592 (5.1)    | 697 (5.4)    | 622 (5.0)    | 588 (4.5)    | 2499 (5.0)    |
| <b>Total</b> | <b>11671</b> | <b>12974</b> | <b>12439</b> | <b>13116</b> | <b>50200</b>  |

**Table 2** Pathology Results for 7332 Women With Breast Biopsies, n (%)

| Age, y       | Invasive cancer    | DCIS             | LCIS            | Atypical hyperplasia <sup>a</sup> | Other atypical types <sup>b</sup> | Intraductal papilloma | Radial Scar     | Sclerosing adenosis | Fibro-adenoma      | Hyperplasia without atypia <sup>c</sup> | Benign non-proliferative |
|--------------|--------------------|------------------|-----------------|-----------------------------------|-----------------------------------|-----------------------|-----------------|---------------------|--------------------|---|--------------------------|
| <40          | 55                 | 11               | 2               | 22                                | 4                                 | 48                    | 7               | 9                   | 450                | 30                                      | 377                      |
| 40-49        | 229                | 94               | 19              | 125                               | 27                                | 70                    | 23              | 26                  | 386                | 98                                      | 726                      |
| 50-59        | 418                | 147              | 40              | 134                               | 38                                | 64                    | 15              | 27                  | 201                | 78                                      | 698                      |
| 60-69        | 476                | 128              | 11              | 93                                | 20                                | 56                    | 4               | 12                  | 159                | 50                                      | 466                      |
| 70-79        | 297                | 69               | 8               | 34                                | 11                                | 25                    | 3               | 6                   | 67                 | 21                                      | 212                      |
| ≥80          | 234                | 42               | 2               | 17                                | 6                                 | 7                     | 3               | 0                   | 19                 | 8                                       | 68                       |
| <b>Total</b> | <b>1709 (23.3)</b> | <b>491 (6.7)</b> | <b>82 (1.1)</b> | <b>425 (6.0)</b>                  | <b>106 (1.4)</b>                  | <b>270 (3.7)</b>      | <b>55 (0.8)</b> | <b>80 (1.1)</b>     | <b>1282 (17.5)</b> | <b>285 (3.9)</b>                        | <b>2547 (34.7)</b>       |

<sup>a</sup> Atypical ductal hyperplasia, atypical lobular hyperplasia.

<sup>b</sup> Flat epithelial atypia, papillary atypia, apocrine atypia.

<sup>c</sup> Columnar hyperplasia, usual ductal hyperplasia.

Abbreviations: DCIS, ductal carcinoma in situ; LCIS, lobular carcinoma in situ.

model), completeness (ie, how well the data variable populates patient records), and correctness (ie, how well the data agree with the most reliable source or gold standard). These domains are important in considering the clinical utility of the clinical data sources and are not addressed by validation alone.

Personalized healthcare is enhanced by improving patient access to data sources both within and outside the healthcare setting. Through interactive systems, patients could have multidirectional interactions with healthcare providers and systems to help make appropriate healthcare decisions that improve their health. To date, this project has not provided direct access to patients, but that work is planned.

Personalized healthcare also requires shared informed decision making between patients and clinicians. While this concept has been widely accepted, the development of practical tools, such as decision aids, to

support shared decision making has only recently emerged. A 2011 Cochrane review on decision aids reported improved patient knowledge (mean difference 13.77 out of 100; 95% confidence interval [CI], 11.40 to 16.15; 26 trials) and risk perception (relative risk [RR] 1.75; 95% CI, 1.46 to 2.08; 14 trials) when probabilities were presented to patients compared to usual care.<sup>21</sup> Also, decision aids that included values clarification improved the proportion of patients who made decisions consistent with their own priorities or preferences.<sup>21</sup>

Decision aids provide women with customized prognostic risk information while engaging them in considering their personal values surrounding the benefits and harms associated with various healthcare options. As a woman begins the decision aid, she uses menus to select key variables to personalize the risk information she receives. Risk information is then presented using various well-accepted approaches includ-

ing graphics (eg, pictographs) and numerical and text explanations. Each appropriate healthcare option is presented based on evidence-based standards of care, and women are able to set priorities for their decisions. Shared decision-making projects have been prioritized by the clinical team and are in development. These include patient decision aids for screening mammography,<sup>22</sup> clinical management after breast biopsy, and breast cancer risk assessment to guide referrals to genetic counseling.

Although this project was conducted in a health system in the United States and focused on a specific clinical condition, its methods can be applied broadly. A clinical team identifying data priorities, defining data elements using a common data model, and creating clinical pathways can be assembled from existing medical staff. A technical team, either in-house or consulting, is required to identify existing data sources or develop them if they are lacking and provide linkages. Each project can be customized to the needs and resources of the health system or institution and modified over time.

The success of this project is based on an incremental approach to extracting relevant data from disparate noninteroperable data systems. The creation of a data warehousing model and a customized data mart is an effective solution. However, true multidirectional information exchange between data sources and the breast health data mart at the point of care, within patient's personal health records, and outside the health system will ultimately provide the most personalized connected data system.

#### REFERENCES

- Lau F, Kuziemsky C, Price M, Gardner J. A review on systematic reviews of health information system studies. *J Am Med Inform Assoc*. 2010;17(6):637-45.
- Buckley JM, Coopey SB, Sharko J, et al. The feasibility of using natural language processing to extract clinical information from breast pathology reports. *J Pathol Inform*. 2010;3:23-29.
- Jaspers MWM, Smeulders M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. *J Am Med Inform Assoc*. 2011;18(3):327-34.
- US Preventive Services Task Force. Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2009;151(10):716-26.
- Smith RA, Saslow D, Sawyer KA, et al. American Cancer Society guidelines for breast cancer screening: Update 2003. *CA Cancer J Clin*. 2003;53(3):141-69.
- National Cancer Institute. Breast Cancer. 2013. <http://www.cancer.gov/cancer-topics/types/breast> on 4 March 2013. Accessed August 15, 2013.
- Ries LAG, Melbert D, Krapcho M, et al. SEER Cancer Statistics Review, 1975-2004. National Cancer Institute. Bethesda, MD; 2007.
- Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: systematic evidence review to update the US Preventive Services Task Force. *Ann Intern Med*. 2009;151(10):727-37.
- Ashbeck EL, Rosenberg RD, Stauber PM, Key CR. Benign breast biopsy diagnosis and subsequent risk of breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2007;16(3):467-72.
- Coopey SB, Mazzola E, Buckley JM, et al. The role of chemoprevention in modifying the risk of breast cancer in women with atypical breast lesions. *Breast Cancer Res Treat*. 2012;136(3):627-33.
- National Cancer Institute. Screening mammograms: questions and answers. <http://www.cancer.gov/cancertopics/factsheet/Detection/screening-mammograms> Accessed July 30, 2012.
- US Preventive Services Task Force. Genetic risk assessment and BRCA mutation testing for breast and ovarian cancer susceptibility: recommendation statement. *Ann Intern Med*. 2005;143(5):355-61.
- US Preventive Services Task Force. Chemoprevention of breast cancer: recommendations and rationale. *Ann Intern Med*. 2002;137(1):56-8.

- Nelson HD, Smith MEB, Griffin JC, Fu R. Use of medications to reduce risk for primary breast cancer: a systematic review for the US Preventive Services Task Force. *Ann Intern Med*. 2013;158(8):604-14.
- Krishnamurthy S, Bevers T, Kuerer H, Yang WT. Multidisciplinary considerations in the management of high-risk breast lesions. *AJR Am J Roentgenol*. 2012;198(2):W132-40.
- Stang PE, Ryan PB, Racoosin JA, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med*. 2010;153(9):600-6.
- Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database. *AJR Am J Roentgenol*. 1997;169(4):1001-8.
- National Quality Measures for Breast Centers. <http://www/nqmbc.org>. Accessed August 15, 2013.
- Oster NV, Carney PA, Allison KH, et al. Development of a diagnostic test set to assess agreement in breast pathology: practical application of the Guidelines for Reporting Reliability and Agreement Studies (GRRAS). *BMC Women's Health*. 2013;13:3. doi:10.1186/1472-6874-13-3.
- Nelson HD, Wang L, Weerasinghe R, Grunkemeier G. Trends and influences on mammography screening in a community health system. 2011 Women's Health Congress. *J Womens Health*. 2011;20(3):474.
- Stacey D, Bennett CL, Barry MJ, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev*. 2011 Oct 5(10):CD001431.
- Scariati P, Nelson LN, Kalpathy-Cramer J, Eden KB. Using a web-based tool to help women make informed choices about breast cancer screening. Proceedings of 2012 NLM Training Conference, Madison, WI; 2012 June 26-27. <http://www.nlm.nih.gov/ep/documents/TC2012TrainingConferenceBook.pdf>. Accessed August 15, 2013.

**Acknowledgments**  
This work was funded by the Providence Cancer Center and the Safeway Foundation.