Li-Fraumeni Exploration Consortium Data Coordinating Center: Building an Interactive Web-Based Resource for Collaborative International Cancer Epidemiology Research for a Rare Condition


ABSTRACT

Background: The success of multisite collaborative research relies on effective data collection, harmonization, and aggregation strategies. Data Coordination Centers (DCC) serve to facilitate the implementation of these strategies. The utility of a DCC can be particularly relevant for research on rare diseases where collaboration from multiple sites to amass large aggregate datasets is essential. However, approaches to building a DCC have been scarcely documented.

Methods: The Li-Fraumeni Exploration (LiFE) Consortium’s DCC was created using multiple open source packages, including LAM/G Application (Linux, Apache, MySQL, Grails), Extraction-Transformation-Loading (ETL) Pentaho Data Integration Tool, and the Saiku-Mondrian client. This document serves as a resource for building a rare disease DCC for multi-institutional collaborative research.

Results: The primary scientific and technological objective to create an online central repository into which data from all participating sites could be deposited, harmonized, aggregated, disseminated, and analyzed was completed. The cohort now include 2,193 participants from six contributing sites, including 1,354 individuals from families with a pathogenic or likely variant in TP53. Data on cancer diagnoses are also available. Challenges and lessons learned are summarized.

Conclusions: The methods leveraged mitigate challenges associated with successfully developing a DCCs technical infrastructure, data harmonization efforts, communications, and software development and applications.

Impact: These methods can serve as a framework in establishing other collaborative research efforts. Data from the consortium will serve as a great resource for collaborative research to improve knowledge on, and the ability to care for, individuals and families with Li-Fraumeni syndrome.

Introduction

Epidemiologic research frequently requires large data collection efforts to establish the natural history of a disease. Collaborative research, especially for rare diseases, is often needed to achieve sufficiently large sample sizes to allow for meaningful study design, and interpretations of findings. However, collaborative research can present technical challenges to researchers with limited access to information technology and bioinformatics resources. Data Coordinating Centers (DCCs) offer a technical and strategic solution that helps mitigate the administrative, computational, and infrastructure burden from individual researchers seeking to conduct multisite collaborative research and enable the implementation of projects based on aggregate data (1).

A DCC is designed, among other functions, to manage the administrative aspects of building a shared central data repository for multiple research sites by implementing effective infrastructures for communication, data collection and data harmonization, and establishment of a central data repository (2). Experience from large epidemiology consortia proves that data harmonization is vital to the development of an aggregated database because submitting institutions often collect and record data differently (3). DCCs must also have the technical expertise to build a highly curated centralized data repository, with embedded statistical analysis tools, and capability for interactive data query and visualization tools.
There is a dearth of published guidance available for the building of a successful DCC for research consortia, especially those involving rare diseases (1). To address this issue, the National Cancer Institute (NCI), Division of Cancer Control and Population Sciences (DCCPS), Epidemiology and Genomics Research Program (EGRP) provided funding for a pilot project to establish a DCC to support the Li-Fraumeni Exploration (LiFE) Consortium. The EGRP supports national and international research consortia focusing on interdisciplinary and translational research of rare cancers. The objective of this project was to expedite collaborative research across Cancer Epidemiology Consortia (CEC) supported by the NCI and other agencies. Specifically, this was to be accomplished through a demonstration project aimed at establishing an efficient and cost-effective DCC for CEC focusing on rare cancers, and establishing guidelines and Standard Operating Procedures for the creation of such DCC to be used by established, emerging, and future rare cancer CECs. The DCC to be created was tasked with providing support for data collection, harmonization, management, distribution, and access to the CEC investigators and, through an approved process, to collaborators external to the CEC.

The LiFE Consortium was created with the mission to foster communication among investigators, and to promote collaborative research projects to advance our understanding of Li-Fraumeni syndrome (LFS) and its impact on affected families. The LiFE Consortium also provides a platform for joint activities between professionals and patients and families to promote support, education, and awareness (4).

LFS (OMIM#151623) is a rare autosomal-dominant cancer predisposition syndrome (5) associated with germline pathogenic variants in the TP53 tumor suppressor gene (6, 7) and is characterized by a high lifetime risk of developing a wide spectrum of childhood and adult onset cancers with osteosarcoma, soft-tissue sarcomas (STS), early-onset breast cancer, brain tumors, leukemia, and adrenal cortical carcinoma (ACCC) being the core cancers (8–11). As more families with TP53 mutations are identified, the LFS cancer spectrum has expanded to include melanoma, lung, gastrointestinal tract, thyroid, ovarian, and other cancers (8, 11–13). Cumulative cancer risk associated with LFS has been estimated to be approximately 50% by age 40 and up to 90% by age 60 (9, 14) with females having higher risk than males (9, 15–17). Clinical diagnostic criteria for classic LFS kindred include a person with sarcoma diagnosed before age 45, with a first-degree relative with any cancer before age 45 and another first- or second-degree relative with a sarcoma at any age or another cancer before age 45 (18). The less stringent Li-Fraumeni-like (LFL) criteria expand the proband’s cancer type to include childhood cancers, brain cancers, and ACC, and extend the relatives’ age at diagnosis to <60 years (19, 20). Germline TP53 pathogenic variants are identified in approximately 70% of families meeting the classic LFS diagnostic criteria (21, 22) and approximately 40% of families meeting the LFL diagnostic criteria (19). The frequency of de novo pathogenic variants in TP53 is estimated to be between 7% and 20% (23). Guidelines for TP53 genetic testing have also been developed (8, 24–28). Although progress has been made because LFS was first described in 1969, many questions related to the natural history of this condition and effective clinical management remain unanswered. Given the rarity of LFS, the assembling of information from multiple research institutions to build a larger dataset is essential in the effort to improve understanding of the condition and patient care.

Materials and Methods

A contract was established with Enterprise Science and Computing (ESAC), Inc., a Biomedical Research Data Management and Health Information Technology company, to build and manage the LiFE DCC. Seven LiFE Consortium participating institutions in the United States, Canada, Brazil, and the United Kingdom were involved in the initial effort to develop a data repository.

Creating the LiFE DCC involved establishing a technical working group and training on the functioning and processes of the repository. A Web-based query tool was also built into the repository for summary data. The LiFE DCC Web portal was designed to house and manage consortium-related public and semi-public information (Fig. 1).

Collaboration and communication infrastructure development

A Steering Committee, consisting of Principal Investigators (PI) from each participating site, Project Leads from ESAC, Inc., and a representative from EGRP, was established at the outset. A Technical Working Group, with research database administrators from each contributing site, was also formed at the beginning of this project. The Steering Committee made decisions regarding all aspects of the development and implementation of the DCC, including data elements, data formatting, inclusion criteria, and prioritization of consortium-related activities. Steering Committee members worked with each other and within their institutions to implement the appropriate agreements, policies, and procedures for data transfer, housing, and use. The Steering Committee also determined the approval process permitting access to data in the LiFE DCC central repository.

The Technical Working Group and ESAC Project Leads carried out the implementation data harmonization, including the data elements collected, data formats, values of each data field, and validation rules for data submitted to the LiFE DCC. The Technical Working Group was also responsible for determining submission methods, creating data dictionaries used to standardize submissions across sites, implementing quality control procedures, and keeping validation logs. Once the technical details of creating the harmonized LiFE DCC central repository were determined, data from each site were prepared in accordance with the established rules, and submitted to ESAC, Inc.

One of the barriers to the development of a DCC is establishing coordinated communications among consortium members and with the DCC (29). The LiFE consortium initially included participants located in Brazil, Canada, the United Kingdom (UK), and the United States, as well as the LiFE DCC Project Leads at ESAC, Inc., and the EGRP representative. The communication barrier has eased with the availability and accessibility of various Web-based communication applications. For the LiFE DCC, communications were facilitated using email correspondence, online meetings, weekly to semi-weekly teleconferences, and Web conferences, and an online customized Wiki space.

In anticipation of the multiple documents and resources to be shared, the LiFE DCC developed a password-protected online Wiki space to organize and provide members with easy access to all consortium-related information and documents. The functionalities of the online Wiki space included sharing and downloading documents, storing files, listing due dates for specific tasks, and posting of meeting minutes and other relevant information. The Wiki also served as a space for LiFE Consortium members to communicate between teleconferences. The Wiki space was vital in organizing all pertinent information and documentation in one convenient and easily accessible online location.
Data submission, harmonization, management, quality control, and assurance

A Data Transfer Agreement (DTA) was drafted by the NCI's Technology Transfer Team to be executed between each participating LiFE Consortium site and ESAC, Inc. The DTA outlined the handling of data by the DCC and the duration of the agreement. Data could only be submitted to the DCC after the DTA between the contributing site and ESAC, Inc. had been executed.

A Data Use Agreement (DUA) was also created by the NCI's Technology Transfer Team to define the responsibilities of the contributing LiFE Consortium sites with respect to data use, sharing, protection, and accreditation. The DUA was executed amongst the participating LiFE sites, allowing each site access to the aggregate harmonized data housed in the central repository.

Data harmonization is a vital undertaking for a successful DCC. The establishment of a common data dictionary is essential to organizing all data within a DCC (30). Each consortium site may have its own format and values for each data element. To facilitate the harmonization of data among the sites, Technical Working Group members submitted a list of all the data elements collected, and the permissible values for each element, contained in their respective databases in a Microsoft Excel (or equivalent) spreadsheet. The DCC then created a table to map and tally all common data elements (CDE) across the submitting sites. The CDE table was compared against existing data standards, namely, the Cancer Data Standards Registry (cDSR), and the United States Health Information Knowledgebase (USHIK). Existing standardized definitions/coding rules for data elements on the CDE list were adopted where possible. Even though similar data were collected, how they were recorded and coded varied significantly from site to site. Especially challenging are the data on cancer diagnosis, as the information was collected in various formats, including text and ICD codes. There was no standard for how the text was recorded. All data submitted for the cancer diagnosis fields were reviewed by a member of the Steering Committee and mapped to a specific diagnosis.

On the basis of the CDE analysis, a proposed Minimum Data Set (MDS) was assembled, and approved by the Steering Committee, consisting of 20 data elements that were common to at least three of the seven participating LiFE sites. An additional 27 optional data elements were considered, with 13 added to the final set of elements to be included in the central repository. Coding and validation rules for each data element were determined by the Technical Working Group (TWG) and recorded in the Data Dictionary. Rules for transforming the values of each data element to conform with the format of the central repository were also established. Test cases from all sites were sent securely via an open-source secure File Transfer Protocol (sFTP) client (FileZilla https://filezilla-project.org/client_features.php) to test the implementation of the LiFE DCC Data Dictionary. Once the fidelity of the transferred test data was assured, all sites with established DTAs submitted all data in accordance with the finalized data dictionary. The sites submitted their data over the sFTP client, or via encrypted email.

Manual quality assurance (QA) and automated quality control (QC) measures were implemented by the DCC to ensure the fidelity of the submitted and transformed data. Manual QA checks included mapping submitted data elements and data formats to the DCC data dictionary equivalents, comparing counts before and after data harmonization and transformation, and confirming the integrity of the data.
transformed into the database from 20 randomly selected cases from each submitted dataset. All datasets were double checked to ensure that no Personal or Protected Health Information (PHI) was inadvertently submitted to, and stored within, the DCC.

The automated QC solution uses a rule-based validation system based on the LiFE DCC Data Dictionary and is implemented via a Pentaho Data Integration tool within the data Extraction, Transformation, and Loading (ETL) process. This facilitates error detection and correction for data submitted. Validation rules comprise both standard checks on allowable values, and a crosscheck of related database elements for logical and scientific consistency (e.g., Vital Status vs. Year of Death). Validation errors were collected and logged within the database when the submitted data could not be harmonized.

The flexible architecture of the LiFE database allows for exceptions to the validation rules as determined by the Steering Committee. Certain data were allowed into the LiFE database despite validation errors. For example, the format for the “Cancer Diagnosis” field was initially built to accept an ICD-9/10 value; however, this information was not readily available for all subjects in the sites initially built to accept an ICD-9/10 value; however, this information was not readily available for all subjects in the sites’ datasets. Thus, the validation rules for this field were revised to accept cancer diagnosis text values as well. This permissible violation allowed for the data to be captured, but raised validation error flags in the QC process.

For any errors or inconsistencies noted, the DCC contacted the respective site and requested a resubmission of the data. All errors and inconsistencies identified were satisfactorily addressed. Each site also received their transformed data after each data load for comparison against their own submitted files for any discrepancies, independent of the LiFE DCC QC and QA procedures.

Data access

To facilitate access to the central repository housed at the LiFE DCC, the DCC built a Web portal consisting of a publicly viewable section, and a user-restricted access component. The publicly viewable content on the Web portal is comprised of information about LFS, the consortium, and real-time summary level descriptive data tables, graphs, and charts are generated and displayed live using the Saiku query and visualization interface (Fig. 2). There are several filtering options available when making queries to generate highly specific sets of data. All datasets and graphics can be downloaded for further analysis. The publicly viewable summary level descriptive data tables, graphs, and charts are generated and displayed live using the Saiku query and visualization interface (Fig. 2).

The technical and software specifications

The LiFE DCC was created using open source packages: LAM/G Application (Linux, Apache, MySQL https://www.mysql.com/, Grails http://grails.org), Extraction-Transformation-Loading (ETL) Pentaho Data Integration Tool (https://www.hitachivantara.com/en-us/products/data-management-analytics/pentaho-platform.html), and the Saiku-Mondrian client (http://community.meteorite.bi, http://mondrian.pentaho.com/documentation/schema.php) for queries made of the central repository. The submitted data went through the ETL process and were loaded into the MySQL Database. The Pentaho Data Integration Tool’s ETL Process (Fig. 4) facilitates harmonization and aggregation of data submitted by the LiFE sites to the LiFE DCC. The extraction step maps all columns from the submitted excel sheets to the LiFE DCC Data Dictionary elements. The data were then loaded into memory for transformation. The transformation step assigned each submitted case with a unique LiFE DCC identification number (ID), mapped all submitted values to the LiFE DCC Data Dictionary, and confirmed there were no validation errors or failed records for each case submitted. Data loading to the central repository is the final step in the automated QC and relational database population process. Manual QA steps are then performed by the DCC before moving the data to production.
The “Reports” page of the public LiFE DCC Web portal provides summary-level data tables and interactive visualization tools. A link to the LiFE DCC Data Dictionary is included as well. The descriptive data tables on display include “Individual Gender Distribution,” “Family Gender Distribution,” and “Biospecimen Availability” by P53 mutation status, a frequency distribution of the “Top 10 Cancer Diagnoses” for individuals affected by LFS, “Age Distribution” by individual and family within the repository, and “Family Recruitment” by country and vital status.

Visual overview of the Extraction, Transformation, and Data Loading (ETL) Process. Files are submitted as excel sheets. The extraction step maps all columns from the submitted excel sheets to the LiFE DCC Data Dictionary elements. The data are then loaded into memory for transformation. The transformation step assigns each submitted case with a unique LiFE DCC ID, maps all submitted values to the LiFE DCC Data Dictionary, and confirms there are no clinical validation errors or failed records for each case submitted. Data loading to the central repository is the final step in the automated QC and relational MySQL database population process.
The Saiku server is an open-source server capable of rendering reports and datasets based on Mondrian schema definition. The Saiku server collects queries from the user, pulls this data from the MySQL database, and reports the results in any format requested by the user including tables, charts, graphs, and other data visualization options (Fig. 5). All reports created in Saiku are rendered in the Web portal. Data are accessed through a RESTful interface running on an Amazon Web Server M3 - large instance.

**Results**

The Life DCC pilot project was completed in July 2016. Table 1 and Fig. 3 show the totals for all cases submitted by each contributing site.

### Table 1. Number of individuals and families included in the LiFE DCC database.

<table>
<thead>
<tr>
<th>Families</th>
<th><strong>TP53 P/LP variants (N = 644)</strong></th>
<th><strong>TP53 VUS (N = 25)</strong></th>
<th><strong>TP53 negative (N = 287)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Carriers</td>
<td>Noncarriers</td>
<td>Untested</td>
</tr>
<tr>
<td>Number of individuals</td>
<td>1354</td>
<td>499</td>
<td>205</td>
</tr>
<tr>
<td>Contributing sites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital A.C. Camargo – Fundação Prudente (ACCC)</td>
<td>208</td>
<td>237</td>
<td>0</td>
</tr>
<tr>
<td>City of Hope (COH)</td>
<td>91</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Dana-Farber Cancer Institute (DFCI)</td>
<td>155</td>
<td>47</td>
<td>0</td>
</tr>
<tr>
<td>Hospital for Sick Children (HSC)</td>
<td></td>
<td>Pending</td>
<td>Pending</td>
</tr>
<tr>
<td>MD Anderson Cancer Center (MDA)</td>
<td>343</td>
<td>146</td>
<td>57</td>
</tr>
<tr>
<td>National Cancer Institute (NCI)</td>
<td>317</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Saint Mary’s Hospital (SMH)</td>
<td>203</td>
<td>46</td>
<td>135</td>
</tr>
<tr>
<td>St. Jude</td>
<td>37</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>Age, mean (range)</td>
<td>35.5 (0–94)</td>
<td>47.3 (6–89)</td>
<td>39.3 (1–90)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>617</td>
<td>104</td>
<td>44</td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>33</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>312</td>
<td>269</td>
<td>8</td>
</tr>
<tr>
<td>Other/not specified</td>
<td>392</td>
<td>114</td>
<td>147</td>
</tr>
<tr>
<td>History of cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>965</td>
<td>54</td>
<td>107</td>
</tr>
<tr>
<td>No</td>
<td>389</td>
<td>445</td>
<td>98</td>
</tr>
</tbody>
</table>

Abbreviations: DCC, Data Coordination Center; P/LP, pathogenic or likely pathogenic; VUS, variant of uncertain significance.
LiFe site and illustrate some examples of the graphics that are publicly available. There are 644 families with a TP53 pathogenic or likely pathogenic (P/LP) variant. Among these families, there are 1,354 individuals who tested positive, ages ranging from 0 to 94 years, 499 who tested negative for the familial P/LP variant, with ages ranging from 6 to 89, and 205 untested individuals, ages ranging from 1 to 90. Approximately 71% (965/1,354) of individuals who tested positive for the familial TP53 P/LP variant had been diagnosed with at least one cancer. Approximately, 37% of the individuals are White, 29% are Hispanic/Latino, and 2.5% are Black. There are also 287 families which meet the classic LFS or LFL diagnostic criteria, but in which no TP53 P/LP variant has been identified. In 25 families, a TP53 variant of uncertain significance (VUS) was identified (Table 1). The clinical data collected included the required and optional data elements. Currently the database contains 48 data elements, including information on contributing centers, demographics, TP53 status, and cancer diagnosis (Fig. 6).

As mandated by the EGRP pilot project, the DCC was established to support the creation of a database and communication infrastructure to facilitate collaborative research. Once these goals were reached, the database was transferred from ESAC, Inc. to City of Hope Comprehensive Cancer Center (COH; a contributing LiFe site) who assumed full responsibility for the management and further development of the resource. ESAC, Inc.’s LiFe DCC Project Leads and Technical Team worked closely with COH’s team members to configure all technical specifications required to host the database and Web portal (http://www.dcc.ac.uk/projects/life), and provided training on all processes required to collect, transform, and upload the data to the database.

Each collaborating site is provided with the data dictionary and associated data entry form, which includes a pedigree-level relational data format, where each person is assigned an ID and is connected to family members via their associated Mother ID and Father ID, as per standard practice for genetic research collaborations. Once the Data Transfer Agreement has been completed between the DCC at COH and the collaborating site, the site can submit their data. Data are verified by DCC staff once received and imported into the database. The collaborating site is provided with access to their own data folder so that they can verify that their data were uploaded as expected, and make updates as needed. Alternatively to sending data to the DCC, the collaborating site can enter the data directly into their own folder.

The DCC is now fully functional and is open to any research group that can contribute genotypic and minimally required phenotypic and epidemiologic data from at least five individuals with LFS with or without a cancer diagnosis. Groups with smaller numbers of participants are encouraged to participate through partnership with a larger group.

Several collaborative studies have resulted from the LiFe consortium, with two publications on the phenotype of TP53-associated breast cancers (31) and findings at baseline for surveillance utilizing rapid sequence whole-body MRI (wb-MRI) among individuals with LFS (32). The LiFe DCC has provided data supporting multiple ongoing projects, including a successful R01 grant proposal (R01 CA242218) attempting to address the NCI Provocative Questions around understanding penetrance and modifiers of inherited TP53 (principal investigators (PI): J.N. Weitzel, J.E. Garber, and C. Amos) for the Li-Fraumeni and TP53: Understanding and Progress (LiFT UP) study (liftup@coh.org). The DCC has also provided potential sample size data for another recently approved concept on a metformin prevention trial under development at the NCI. More sites have joined since the creation of the database. As intended, data from LiFe centralized repository database have been utilized by researchers both internal and external to the consortium. One example of the research being carried out by external investigators using data from the LiFe DCC is a project examining the phenotype for LFS-associated breast cancers. The data is being used to validate previous observations that Her2/neu-amplified breast cancer was seen more frequently among TP53 carriers, and then incorporate phenotype into a multifactorial variant classification algorithm. The LiFe consortium was able to contribute data on 246 TP53-associated breast cancers, and a manuscript is under review. A better understanding of the phenotype of the

Figure 6.
Database Star Schema of core data points collected, by category, including: demographic, family, TP53 mutation details, and diagnoses.
syndrome has the potential to improve the clinical indications for testing and risk management.

Discussion

Establishing a high-quality database is an essential undertaking for any successful collaborative research. This is especially true in rare diseases where data are collected by multiple centers over many years, and the types and methods of data collection might change over time. A DCC can help facilitate data aggregation and harmonization, and ensure the quality of the database created. Here we described the administrative and technical practices of the LiFE DCC, which worked to establish a centralized database. This project met the goal of the EGRP to establish a DCC for a CEC focusing on rare cancers. The operating procedures outlined in this article could be utilized by rare cancer CEC for the establishment of a similar DCC.

In addition to the challenges with the coordination of communication and the creation of the data dictionary discussed above, other challenges encountered during the building of the database and, subsequently, its maintenance, include international data transfer agreements and the ability to add data fields afterwards. International institutions might have different requirements, which would need to be considered. Furthermore, the NCI funded the creation of the database and DCC, but once that was accomplished, a member institution (in this case, COH) needed to assume responsibility for maintenance, and small grants from the patient advocacy group LiFS Association helped support a partial effort coordinator. Furthermore, it was determined by the steering committee that a pedigree-driven relational database would be most helpful, thus Progeny (Delray Beach, FL) was incorporated into the database.

During the course of creating the DCC and the subsequent planning for utilization of the data, several lessons were learned. First, various contributing sites bring different perspectives in terms of data collection. It is essential to define all fields in the data dictionary, and the functionality of the database must be flexible to adapt to evolving permissible values for each field, as well as to new data fields. It is also critical to put in place early in the process plans for continual funding and protocols for data requests and data use.

A goal of a centralized database is ease of access. The LiFE database that was created allowed individual PIs to conduct analyses on available data, including the ability to query records and create charts and figures. No previous coding experience or programming knowledge is necessary, and Progeny helps facilitate data query. Similarly, many clinical cancer genetics programs use the Progeny interface, which helps enhance the capabilities of the data repository.

This query-capable multisite collaborative central data repository serves as a valuable tool for LFS-related research. The Web-based application of data analysis and visualization tools allows for the ability to access aggregate data from anywhere. The LiFE database will help to advance our collaborative research efforts in LFS, and ultimately help advance the care of individuals and families burdened by this condition.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors’ Contributions


Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P.L. Mai, S.R. Sand, N. Saha, L. DiGianni, E.J. Root, R.C. Bremer, J. Bojadzieva, A. Novokmet, K.E. Nichols, J.D. Schiffman, D.G. Evans, M.I. Achatz, L.C. Strong, J.E. Garber, S.A. Ladwa, J.N. Weitzel

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P.L. Mai, N. Saha, D.G. Evans, L.C. Strong, M. Oberti, D. Malkin, J.N. Weitzel


Study supervision: P.L. Mai, L. DiGianni, S.A. Ladwa, J.N. Weitzel

Acknowledgments

This work was supported by the National Cancer Institute at the NIH, Department of Health and Human Services, under contract no. HHSN261201500214P. The Li-Fraumeni Syndrome Association provided partial support to City of Hope for maintaining the LiFE DCC. Research reported in this publication was also supported by the NCI of the NIH under award number R01CA242218. D.G. Evans is supported by the National Institute for Health Research (NIHR) BRC Manchester (grant reference number 1215-200074). The Li-Fraumeni Syndrome Association provided partial funding for this work.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 14, 2019; revised January 9, 2020; accepted March 3, 2020.

References


Li-Fraumeni Exploration Consortium Data Coordinating Center: Building an Interactive Web-Based Resource for Collaborative International Cancer Epidemiology Research for a Rare Condition


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-19-1113

Cited articles
This article cites 31 articles, 13 of which you can access for free at:
http://cebp.aacrjournals.org/content/29/5/927.full#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link http://cebp.aacrjournals.org/content/29/5/927.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.