



**Headquarters**  
185 Berry Street  
Suite 2000  
San Francisco, CA 94107

650.854.9400

**Washington Office &  
Barbara Jordan  
Conference Center**  
1330 G Street NW  
Washington, DC 20005

202.347.5270

May 8, 2024

Chiquita Brooks-LaSure, M.P.P.  
Centers for Medicare & Medicaid Services  
7500 Security Boulevard  
Baltimore, MD 21244  
[VRDCRFI@cms.hhs.gov](mailto:VRDCRFI@cms.hhs.gov)

Dear Administrator Brooks-LaSure,

KFF is pleased to respond to the Centers for Medicare & Medicaid Services (CMS) [request for information](#) about research data requests and access policy changes. KFF currently accesses several CMS Research-Identifiable File (RIF) datasets through physical data extracts and welcomes the opportunity to share feedback in hopes that our and other organizations' eventual transition to the virtual research data center (VRDC) is as seamless as possible.

KFF is a nonprofit health policy research, polling, and news organization. We conduct and communicate research and information about health policy at no charge. We operate with a diverse mix of funding primarily from the resources we earn from managing our endowment, supplemented by funds we receive from foundations and other external funders. Policymakers, the media, and academics have come to rely on KFF as a trusted source of timely national and state-level information and data on health policy.

We are hopeful that the VRDC could offer our organization benefits in the form of faster processing times and prompter access to the most recent CMS data. However, we are concerned that without a revised approach to data access through the VRDC that allows for greater staff collaboration, easier and faster access to output, and a substantially modified fee structure that keeps our costs for virtual data access on par with our existing costs to access physical data extracts, it will be difficult for KFF to maintain our existing staff and research portfolio. We believe that conducting all of our CMS RIF data research within the VRDC environment, as currently proposed, could jeopardize our ability to provide timely research to inform current health policy discussions at both the national and state levels.

Below, we respond to the RFI questions and raise additional questions we have related to the VRDC environment and the transition. In Table 1, we summarize our most significant concerns and suggestions for how to address these concerns assuming all researchers are required to transition to the VRDC. Although most of our comments below focus on the transition to the VRDC, we respectfully suggest that CMS consider a model for data access whereby certain organizations could apply to be a "trusted entity" that could continue to access physical data extracts using current processes after a thorough review of each organization's compliance with data security and privacy practices as outlined in their DMP SAQ. For organizations seeking to be approved as "trusted entities" CMS could also consider establishing additional requirements related to data privacy, security, and other factors related to the public trust.

**Table 1. KFF’s Key Concerns with Transitioning to the Virtual Research Data Center (VRDC)**

Concern	Explanation	Options to Consider
<p>Moving to the VRDC will significantly increase the costs of doing research.</p>	<p>The fees to store, process, and access data and output exceed our current costs for physical data and are high compared to other data custodians such as HCCI.</p>	<ul style="list-style-type: none"> <li>• Cap the costs per DUA to a fixed amount that would cover the cost of storing, processing, and outputting data to ensure greater predictability in VRDC costs. Fixed costs could be set using information on existing costs per DUA at a few different levels for different-sized organizations with different numbers of data users and DUAs of various scopes and sizes.</li> </ul>
<p>The VRDC pricing per “seat” for each user discourages collaboration and training of junior staff.</p>	<p>The only way to make research cost effective under a per-user fee structure is to consolidate programming tasks within a small number of experienced people, which will limit opportunities for training new users.</p>	<ul style="list-style-type: none"> <li>• Determine pricing per DUA rather than per user.</li> <li>• Allow for institutional or shared seats.</li> </ul>
<p>VRDC costs to store and process data are high and hard to estimate.</p>	<p>It’s unclear whether we are required to pay to store the raw data and how the costs of storing code and analytic files will add up in the VRDC, and how quickly we can add storage and processing capacity if we hit the limits that have been purchased at the outset of the project.</p>	<ul style="list-style-type: none"> <li>• Identify ample, cheap storage space—potentially through a partnership with cloud-based services before asking everyone to transition.</li> <li>• Provide more technical assistance to data users in estimating their storage needs.</li> <li>• Cap the costs per DUA to a fixed amount that would cover the cost of storing, processing, and outputting data to ensure greater predictability in VRDC costs per project.</li> </ul>
<p>The output review process limits the timeliness of our analysis and imposes extra costs.</p>	<p>Much of our work seeks to inform policymakers and the media who often want quick answers. We also provide much of our data to researchers and policymakers through detailed state-level tables, charts, and interactives. In the VRDC environment, this will increase the volume of our output reviews and associated costs, and it will make timely research difficult and, at times, impossible.</p>	<ul style="list-style-type: none"> <li>• Establish a streamlined output review process for qualifying organizations or DUAs.</li> <li>• Develop a “fast lane” output review process for non-sensitive output that is needed for timely research related to current policymaking (e.g., responses to RFIs, comments on proposed rulemaking, Congressional requests, pending legislation).</li> </ul>

## **Section 1: CCW VRDC Processes/Access**

- 1. How much lead time will you need to transition your research study into the CCW VRDC? Please include details about the steps you will take, and the anticipated timeframes associated with each step.*

At this time, we lack sufficient information from CMS to outline the steps we will take in our transition process, but we anticipate needing significant time (on the order of 24-36 months) to make a complete transition. Some questions we have about the transition process include the following:

- Will we need to submit new data use agreements (DUAs) through ResDAC for each project that is required to transition into the VRDC or will our existing DUAs be transitioned into the VRDC automatically? The current data request process through ResDAC is extremely time-consuming and if we needed to submit entirely new DUAs, that would add at least 12 months to the above-mentioned time frame we would need to transition to the VRDC.
  - For people within the organization who currently access data on multiple DUAs, will we need to pay separate user access fees for each DUA they work on?
  - Can we expect to receive detailed information about the new costs associated with our current DUAs in the VRDC *prior* to making the transition, and if so, when can we expect to receive this information?
  - How can we estimate our initial Databrick unit consumption for each DUA?
  - How can we estimate our initial data storage needs for each DUA?
  - How can we estimate our needs with respect to “Analytic Containers”?
  - When our existing DUAs transition to the VRDC, will we be required to pay the \$13,000 project renewal fee (based on “full VRDC” pricing) or the \$18,000 new project fee? Since our DUAs are not new, we believe paying the lower project renewal fee would be appropriate.
- 2. What hurdles or challenges do you anticipate you will have with working in the CCW VRDC?*

A key concern we have with working in the VRDC relates to the per-person user access structure and fee schedule. Under the current fee structure, KFF would likely need to make personnel changes involving a reduction in analytic staffing levels, scale back our research output, or secure new funding. KFF investigated moving to the VRDC in prior years but determined it was not the best choice given our collaborative team environment, where several team members are able to work directly with the physical data extracts but are not dedicated programmers on a full-time basis. Based on our current project staffing and the current VRDC fee structure, the per-user costs for each staff person who currently works on projects under our separate DUAs would amount to additional total costs each year in the low to mid six figures, by our best approximation. The per-person nature of the costs means we would likely not be able to support as many skilled programmers in the VRDC environment and they in turn would not be able to

provide as much mentorship and support in that environment to junior staff who are eager to gain programming and research skills.

With the possibility of changes to the VRDC data access policies and fee structure, we would like to suggest for CMS's consideration the following changes to allow organizations that do collaborative work to maintain existing research levels in the new environment without incurring substantial additional costs.

**We offer the following ideas related to data access and pricing for CMS's consideration:**

- Cap the costs per DUA or per organization to a fixed amount that would cover the cost of storing, processing, and outputting data to ensure greater predictability in VRDC costs per project. Fixed costs could be set at a few different levels for different-sized organizations with different numbers of data users and could be set to ensure total costs are on par with the costs organizations currently pay to access physical data extracts. This is not to suggest that costs should be locked in for specific organizations at current levels even if the scope of their DUAs expand or that costs cannot increase with overall inflation in future years, but that it should not be substantially more costly for a given organization to access the same amount of data for a given DUA through the VRDC than through physical data extracts.
- Determine pricing per DUA rather than per user.
- Establish an institutional pricing arrangement that would enable multiple people within an organization to share seats and for the prompt re-assignment of seats when individuals leave the organization or project.
- Establish multiple tiers of users and associated costs, allowing some users to review code and output within the VRDC but not to access the most sensitive data.
- For trusted entities, allow people in the VRDC to extract and output of de-identified person-level files to be used internally by other staff within the organization who do not have access to the VRDC.

Another concern with working in the VRDC relates to the output review process, limits, and timing. One of KFF's central contributions to the policy community is the production and dissemination of [state-level data](#) that adhere to CMS's cell suppression policy, but that researchers and policymakers can easily access for free. Such data increase the volume of output for KFF but provide significant value to the research community. In addition, KFF frequently fields requests from national and local media, Congressional staff, and executive agencies. Our ability to quickly respond with relevant data points increases our ability to be an effective resource to those parties. We are concerned that the 48-hour timeframe for output review and the 1 GB limit on output review size could limit our ability to be responsive to timely policy requests. We are also concerned that, with the volume of output reviews increasing when all CMS RIF data research projects move to the VRDC, it could be difficult for CMS to ensure that even this 48-hour timeframe is met.

**We offer the following ideas related to output review for CMS's consideration:**

- Develop criteria that allow VRDC users to output non-sensitive data without an output review.
- Establish a mechanism for rapid turnaround output review (within 24 hours or less) when a pressing media or policy question demands a timely response.
- Consider increasing the 1 GB limit per week per DUA on output reviews and increasing the number of output reviews for researchers from 3 to 5 or 10.
- Streamline output reviews and associated fees for organizations and types of output that meet certain requirements.

Another source of concern relates to the limited allocation of storage space that comes with the VRDC and the potential costs we might incur related to the need for additional data storage and data processing. KFF has made significant financial investments in our data infrastructure in the past 5 years to accommodate working with multiple years of large RIF datasets and running complex code. We currently have the flexibility to add storage quickly if necessary and with minimal disruption to our workflow. We are concerned about the potential for lengthy disruptions in our work if we run out of data storage or Databricks capacity in between DUA renewals and need to cease work in the VRDC for the time it takes to pay for and add more storage and processing capacity. In addition, for data that have been approved for use under multiple DUAs, we are also concerned about the duplicative charges we would incur to store the same data under each DUA in the VRDC. (Our understanding is that we must pay to store the raw data files for each DUA. If this is not correct, this concern would be eliminated.)

3. *Is there specific training or assistance you will need to be successful in the CCW VRDC? If possible, please indicate the level of training needed and on which tools.*

We are not aware of any tools that would require training in the VRDC but are interested in learning more from CMS about the VRDC environment, the tools that are available, and the current training opportunities for users to work successfully in the VRDC environment.

4. *Would you consider moving your research study to the CCW VRDC prior to the implementation of the new CMS policies? If so, why and when?*

We are not considering moving any of our existing DUAs to the VRDC prior to the requirement that we do so in the future. One of our DUAs involves reuse of data under another DUA, and it is our understanding that we cannot link data accessed in the VRDC to a physical data extract. Therefore, we would need to amend the DUA to include the reuse data as new use, or initiate a new DUA altogether, which takes a significant amount of time to clear through ResDAC and the CMS Privacy Board. For example, we have been working on a new RIF DUA for more than 9 months and still have not finished the ResDAC review process. Given the complicated nature of the data request process, as well as the high level of uncertainty around the new costs we would incur to have all project staff working in the VRDC, we are not interested in moving our existing DUAs into the VRDC until this transition is required.

5. *Are there research studies that you expect to complete in 2024 or 2025? If yes, please provide the Data Use Agreement (DUA) number and expected completion date.*

We do not anticipate completing any CMS RIF DUAs in 2024 or 2025. Much of our research involves providing the most up-to-date information possible to policymakers, the media, and the broader health policy community. As such, we update our work on a regular basis and have benefited from having ongoing access to the new years of CMS data in order to provide timely information and analysis.

6. *How many seats/users do you anticipate having on your research study once transitioned into the CCW VRDC?*
  - a. *Do you anticipate using the Statistical Analysis Software (SAS) only or the full VRDC option? For information on the CCW VRDC access options, please review the About the VRDC and Requesting Access page identified in the announcement.*
  - b. *Will your research study require the purchase of additional space or Databricks credits?*

Our organization currently has three DUAs that would need to be transitioned to the VRDC and is applying for a fourth. We anticipate needing 15 – 20 VRDC seats if people listed in multiple DUAs need multiple seats, at least 25 TB of storage for raw data, and at least 50 TB of storage for total data. We would want the full VRDC option for each DUA. We would appreciate assistance from CMS in determining how this information translates into the purchase of additional data storage and Databricks credits and are happy to share additional details on each DUA if that would be helpful.

## **Section 2: CCW VRDC Tools**

1. *What analytic tools, program languages, or specific analytic packages and libraries are you using for your research study?*
  - a. *Are you using any analytic tools that are not currently available in the CCW VRDC? For information on the analytic tools currently available in the CCW VRDC, please review the About the VRDC and Requesting Access page identified in the announcement.*

We are using SAS, R, and Stata. Our understanding is that all three programs are available in the full VRDC. Within R, we primarily use the data.table, snow, haven, tidyverse, and openxlsx packages. Our analysts also use R markdown for documentation and output purposes.

KFF has concerns that the VRDC may not have the latest software tools available and that it is hard to add software and work with open-access software. We hope that the VRDC environment will consider some of the practices used in other data enclaves. For example, Optum only provides SQL and R in the data enclave but allows users to buy their own software and load it up so that users are able to pay for and use the software they want. We are also interested in whether VRDC users will be able to install R and Stata packages as needed moving forward. Beyond software, we would like access to version-control tools such as Github.

**We suggest the following ideas for CMS’s consideration:**

- Ensure the VRDC environment has sufficient processing power and storage space and the most recent software tools available prior to requiring organizations to transition into it.
  - Enable organizations to use open-access software and install packages as needed for their data analysis.
2. *If you are using analytic tools not currently available in the CCW VRDC, please describe the workstation used to perform research (Central Processing Unit (CPU), memory, Operating System (OS), number of workstations, etc.).*

This question is not applicable to our organization.

**Section 3: Data/Project**

1. *Do you have data files that will need to be uploaded into the CCW VRDC to complete your research? If so, please describe the data and provide details about the files (format, size, etc.).*

We would need to upload the following data into the VRDC (estimated size is 1TB):

- Area Health Resource File,
- SNP comprehensive report files,
- Geographic crosswalks,
- SSA-FIPS codes,
- Urban influence codes,
- Plan benefit package files,
- Medicare Advantage and Part D Landscape files,
- Nursing home compare data,
- Census data, and
- T-MSIS DQ Atlas quality measures.

2. *Do you have project-specific code that will need to be loaded to your CCW VRDC workspace? If so, please describe and provide details about the code (format, size, volume, language, etc.).*

We would need to upload SAS, R, and Stata code into the VRDC. We estimate that we have about 5 GBs of code.

3. Please estimate the amount of data storage growth per year for your DUA, including the total size of current data in your environment and amount of data imported and generated each month.

We anticipate needing at least 25 TB of storage for raw data, and at least 50 TB of storage for total data, with about 10%-15% growth in storage needs each year.

4. *How long does your data need to be retained for your research?*

None of our projects have specific end dates. As noted above, we frequently update analysis that has been previously conducted as soon as new data become available. Currently, we retain physical data extracts for all years of data for which we have received approval. Our preference is to retain code indefinitely, while output and intermediate data files are generally retained for two years.

**Section 4: Data Access Fees**

1. *How does your organization currently cover costs related to IT infrastructure, security, software licensing, etc. when physically receiving CMS data, and what are estimates of these costs? What is the scope of anticipated cost savings your organization could realize by not paying for IT infrastructure, security, software licensing, etc. related to maintaining a physical copy of CMS data?*

Most of our IT infrastructure and associated costs are covered through KFF's operating budget. We typically do not rely on external funding to support ongoing or new expenses in this area. We do not anticipate significant savings from moving to the VRDC—in part, because we will need to maintain our IT infrastructure for the many research projects that use non-RIF CMS data and non-CMS data. We anticipate saving roughly \$10,000 per year on a SAS renewal and possibly an additional several thousand dollars because we may be able to downsize the space allocations on the secure drives where we currently store CMS RIF data.

2. *How many people are currently associated with your research project that require access to record-level (i.e., non-aggregated) data? How do you anticipate the new policies would affect the number and team structure of researchers accessing record-level data for your project, and what would be the impact of any changes? Could some members of your project team contribute at the same level by reviewing aggregated output?*

As noted above, KFF currently has 3 approved RIF DUAs and a fourth under review. Across those 4 DUAs, we would need 17 seats in the VRDC for individuals who currently or will in the future access record-level data via our physical data extracts if cost were not a barrier and assuming individuals who are listed on multiple DUAs require multiple seats. If VRDC access policy were to allow a single user to access multiple DUAs under one paid seat for that individual, we would have need 14 seats.

If the new policies require KFF to access all data through the VRDC with the current pricing model, we may need to change the composition of our analytic teams somewhat, if not substantially, and potentially change our overall approach to staffing our research using these CMS RIF data. Under our current model, most staff that analyze the data generally also contribute to writing and qualitative analyses that allow them to grow into subject matter experts. Of the 14 individuals who would be accessing the VRDC (which would correspond to 17 separate seats), only 3 are designated full-time programmers. KFF uses this staffing model because it:



- Allows staff to grow professionally within the organization, developing both analytical and subject matter expertise,
- Results in higher-quality analysis because people with deep program knowledge and training in economics/statistics are working directly with the data, and
- Improves the rigor of the analyses by allowing analysts to help each other troubleshoot code, replicate analyses, and conduct rigorous code review by a second expert.

If there are no changes to the VRDC pricing structure, KFF may need to move to a model that segments research staff into “programmer” and “non-programmer” roles. People with seats in the VRDC would be expected to spend most of their working hours in the VRDC, while staff that do not have VRDC seats would review output and do most of the writing and qualitative work. Under this model, we might anticipate higher rates of staff turnover, particularly among programmers, and if we had less continuity in this role, the quality of the analysis might suffer.

3. *Could other types of lower-cost CCW VRDC access meet your needs (e.g., a viewer role that doesn't have access to any analytic tools or software)? If so, what types of roles would you need?*

At this time, we expect most if not all users included in our user counts above would need access to the full VRDC, including analytic tools and software. If there were additional roles that allowed for viewer access without use of analytic tools or software, and the user fee structure was substantially lower, we might in fact consider purchasing additional seats for a handful of additional subject matter experts who work with data but do not need access to record-level data and do not themselves do any programming.

If there was a low-cost, reviewer-only option, we might also consider purchasing additional seats for senior staff who do not work with the data but do review output to help guide analytic decision-making. The purpose of those seats would be so that managers could review output in the VRDC without sending output through a formal CMS output review, enabling faster iteration on coding updates and the output process.

4. *How do your anticipated CCW VRDC fees compare to the total data access fees and internal IT costs associated with your research project?*

We have a substantial degree of uncertainty regarding the potential costs associated with storing and processing data within the VRDC and paying for output, but if the current VRDC fee structure remains intact, we would need to budget for an annual increase in spending on data access fees in the mid six figures, after accounting for very modest savings associated with a reduction in our own data storage space and not renewing our SAS license.

5. *CMS is required to recoup the cost of making data available to researchers to allow the agency to continue offering this important service. Do you have suggestions for an alternative fee structure that would allow CMS to recoup fees associated with VRDC use?*

**We recognize the need to recoup the costs of making data available to researchers and offer the following ideas for CMS’s consideration:**

- Establish a process for institutions that meet defined data privacy and security standards to apply to be “trusted entities,” who can continue accessing the data using similar processes as today.
  - Determine prices per DUA or establish a cap on costs per DUA to a fixed amount that would cover the cost of storing, processing, and outputting data to ensure greater predictability in VRDC costs. Fixed costs could be set at a few different levels for different-sized organizations with different numbers of data users.
  - Allow for institutional access to the VRDC instead of pricing per user.
  - Establish multiple tiers of users and associated costs, allowing some people to review code and output but not to access the most sensitive data.
  - Allow for the extraction and output of de-identified person-level files to be used internally by researchers within an organization who do not have access to the VRDC.
  - Streamline output reviews and associated fees for organizations and types of output that meet certain requirements.
  - Ensure there are no duplicative costs for organizations with multiple DUAs (e.g., requiring organizations to purchase two seats for the same user who is listed on two DUAs; requiring organizations to pay to store the same data being used under separate projects).
6. *How many student dissertation projects does your organization expect to conduct on an annual basis? Based on use, do you have suggestions for the fee structure for dissertation projects?*

We do not routinely conduct or support the completion of student dissertation projects at KFF using our CMS RIF data. The results of our CMS RIF data research projects are disseminated in briefs, data notes, state-level tables, and other public-facing products on our website and other channels of communication.

We do sometimes bring on university-affiliated research fellows who undertake specific research projects during their tenure at KFF. Historically, these individuals have not worked with CMS RIF data, but in theory, they could access and analyze the physical data extracts if needed for specific research projects, subject to the same data privacy and security standards as other KFF staff. Moving to the VRDC, however, would prevent these individuals from accessing data directly since the funds are not available for these individuals to pay for a new seat in the VRDC. CMS could consider creating a lower-cost (or even no-cost) pathway to VRDC access for specific individuals who are temporarily affiliated with an organization with a RIF DUA to conduct

discrete projects on a time-limited basis. Such a pathway could help support dissertation research conducted by students in university settings.

7. *Would it be valuable for CMS to expand the dissemination of lower-cost limited data sets that would not require VRDC access to promote more training and research opportunities for students and other researchers?*

Yes, we believe this would be valuable because it would increase the number of people who are able to work with CMS data. This may help defray some of our costs of moving to the VRDC to the extent that some analysts could work with limited data sets instead of the full record-level data. Along these same lines, perhaps CMS could also consider making LDS data available to researchers who have DUAs for RIF data for a nominal fee, so that when researchers go through the RIF DUA request process, they could request approval for the LDS data at the same time – rather than having to conduct two separate requests.

KFF would welcome additional LDS files but would continue using RIF data for much of our ongoing research. In the Medicaid space, the T-MSIS data do not have an LDS version and the Medicare LDS claims files only provide 5% samples for some services (vs. the 20% samples that we obtain as RIF data). If CMS does plan to enhance the number and type of data files available as LDS files, larger sample sizes (such as 20%) would increase the range of circumstances in which the LDS files would be sufficient for research purposes. If the LDS versions were large enough to support state-level analysis, that would also increase the number of situations in which we could use them, since we aim to publish state-level data and analysis whenever possible as a tool for policymakers and other researchers.

## **Section 5: Transition Timing**

1. *CMS announced plans to require all new RIF Data Use Agreement (DUA) requests to access RIF data within the CCW VRDC beginning on August 19, 2024. If 6 months of advance notice about this change is not sufficient, how much notice would allow you to prepare for this transition? In the interim, what additional security assessments and conditions would you prioritize to address growing security and privacy risks?*

We believe that a period of 24-36 months would provide us with sufficient time to prepare for the VRDC transition, if not longer. This period assumes the VRDC cost structure would be largely unchanged from how it currently operates, which will result in significant new costs that our organization must address. If the cost structure changes, we might be able to transition more quickly.

Assuming no major changes to the cost structure, transitioning will require taking the following actions:

- We will need to re-evaluate project staffing for CMS data projects, including possibly hiring new programmers and transitioning existing staff into other roles where possible.

- KFF currently pays for data storage out of KFF’s operating budget, but some purchases of physical data are funded through external grants that are on two-year grant cycles. The new data storage and processing costs per DUA in the VRDC will exceed what we can pay for out of our operating budget, so we will need to build those costs into the external grants or scale back those research projects. With a two-year grant cycle, we need sufficient time to budget for higher data costs in grant renewals.
- Under the current VRDC cost structure, we also may need to re-evaluate our research portfolio to identify which projects are the biggest priorities, as it may be difficult to sustain our current scope of work with the added expenses.

As far as what additional security assessments and conditions we could prioritize in the interim to address growing security and privacy risks, we could commit to taking further steps to conduct routine (annual or semi-annual) data security and privacy best practices training sessions internally with all CMS RIF data users. Moreover, we have an active Data Management Plan Self-Attestation Questionnaire, originally approved in 2023 and recently recertified in March 2024, which governs our data environment and data privacy and security practices. We would appreciate receiving more information from CMS about what steps organizations could take to reinforce data security and privacy best practices that are already required under the DMP SAQ. We would also suggest for CMS’s consideration that CMS or its VRDC contractor develop, support, or sanction some form of annual online training program on data security and privacy that all organizations with RIF DUAs and data users within those organizations would be required to conduct, including organizations that receive physical data extracts.

*2. To cover growing costs associated with physical data delivery, CMS is updating fees for physical delivery of CMS data beginning on August 19, 2024. If 6 months of advance notice about this change is not sufficient, how much notice would allow you to prepare for the updated data fees?*

Because some of our data purchases are funded through external grants, we would appreciate at least 24 months advance notice to build the new costs into updated contracts. Additionally, we would like information on whether we would be required to pay the \$10,000 project renewal fee for our existing DUAs upon their renewal, or whether we would be charged the higher \$20,000 “initial project fee.” Since our DUAs are not new, we believe paying the lower project renewal fee would be more appropriate.

In addition, we would like to suggest for CMS’s consideration whether it might be more secure, efficient, and cost effective for CMS to provide physical data extracts via a secure FTP (file transfer protocol) mechanism. This could go a long way towards ensuring greater security in the provision of physical data extracts, both in the time before transition to the VRDC and also in the context of the proposed model whereby “trusted entities” continue to have access to physical data extracts.

3. *What other factors not addressed above should CMS consider in determining transition timing for phase 1 or phase 2?*

Some other concerns and questions include the following:

- How will CMS ensure that current RIF data users do not experience any delays in working on projects that are in progress with physical data extracts during the transition? We would request to be granted a window of time (ideally 6-12 months minimum) when we have access to our physical data extracts as well as data in the VRDC to give staff time to get up to speed in the VRDC environment, upload code and any necessary files, and conduct other tasks needed to work successfully with data in the VRDC.
- When will the final fee structures be announced? For our organization to incur hundreds of thousands of dollars in new costs each year, more transition time would be needed. If the cost structure is amended along the lines of the suggestions we have submitted, we may be able to make the transition in less time.
- Does the VRDC have the necessary computing infrastructure to accommodate the influx of research and data storage needs for all RIF DUAs that will be transitioning to the VRDC? We are concerned about the possibility of VRDC outages and/or slow processing times that would disrupt our work and the work of other researchers.
- Does CMS's VRDC contractor have sufficient staff to accommodate the increased demand for output reviews? How will CMS ensure that the added volume of output reviews does not overwhelm the capacity of the system to respond within the 48-hour stated timeframe? And how are output reviews triaged – is it first come, first served, or is there some other mechanism for processing these requests?
- With the influx of new VRDC users, how will CMS ensure that the necessary onboarding and training is conducted expeditiously, and is there a plan for enhanced technical support to address questions and problems that arise in the VRDC environment for both new and existing users?