CareSet Journal: FOIA Domain Name Analysis Whitepaper

Introduction

Do patients benefit from the money that the U.S. Government has spent on the interoperability of clinical systems? The Centers for Medicare and Medicaid Services (CMS) believes that the purpose of the $34 billion dollar Electronic Health Record (EHR) Incentive Programs (aka "Meaningful Use"), is to provide patients with access to their medical records in digital form. CMS even renamed this to the "Promoting Interoperability (PI) Programs". This is the explanation for that change:

> *CMS is renaming the EHR Incentive Programs to the Promoting Interoperability (PI) Programs to continue the agency's focus on improving patients' access to health information and reducing the time and cost required of providers to comply with the programs' requirements.* (emphasis ours)

This language, which is emblematic of other statements from CMS, indicates that CMS regards patients having online access to their own medical records as its statutory duty. Indeed, multiple laws and funding measures passed by Congress give HHS and CMS the duty to ensure that patients, especially Medicare and Medicaid beneficiaries, have access to their clinical data.

CMS also has a statutory duty to report on information blocking. Information blocking rules and HIPAA require that patients are able to get their medical records in the manner they request if the provider has the technical ability to do so (unless exceptions or grounds for denial are met). Providers participating in the Promoting Interoperability Program have EHR technology that give providers the ability to provide a variety of methods of online access for patients. HHS provided an example of how this can work in their guidance explaining the intersection of HIPAA and Information Blocking:

> *For example, in exercising her right of access under the HIPAA Privacy Rule, an individual could request a copy of her information that constitutes the CCDS* [Common Clinical Data Set] *through the provider's Certified EHR Technology portal or that it be sent from the Certified EHR Technology to the individual's Direct address (an electronic address for securely exchanging health information using the Direct technical standard). If the provider is using Certified EHR Technology, the HIPAA Privacy Rule requires the provider to grant this request from the individual because the form and format requested is "readily producible" using the provider's Certified EHR Technology.*

Using FOIA, we have obtained records to shed light on CMS's performance of their duties. CMS states that these incentive payments were intended to ensure that patients have access to their own healthcare data online. But have these payments resulted in healthcare providers actually making this online access easy for patients? What portion of these incentives payments are waste?

Put another way: The US government paid healthcare providers to stop faxing. Did providers do that? Or are they still asking patients to deal with fax machines to get access to their data?

The results of our FOIA request show that many healthcare providers are doing the right thing, and are allowing patients to download their healthcare records easily online. However, this trend is hardly universal, and the number of patients who still do not apparently have non-fax, non-mail, non-in-person, online access to digital versions of their medical records is significant. In our sample, 18 percent did not have a convenient way to use standard medical record request forms to request online access.

Our sample only included a small fraction of the relevant providers (we got less than 3% of the data that we requested) and of these we only modeled a portion of the data (details below) but we still identified about $150 million dollars that had been given to providers who only mention fax, snail-mail or in-person methods of access on their default HIPAA medical record request forms. Our most conservative estimate is that this has an impact on at least 9 million patients.

FOIA Background

This data is the result of a FOIA request made to CMS by journalists at CareSet in 2014.

We had originally requested the full email addresses of providers in NPPES, but we modified our request to just domain names. After waiting for more than five years, the CMS FOIA office told us that in fact we did not have an agreement like we thought, and that we would not be getting this data because of privacy and cybersecurity concerns.

We took CMS to court and won the right to some of the data that we requested.

The court's reasoning is that once a given provider published their domain name under the endpoints PUF file, they have been exposed to any downsides of releasing a domain name to the public. Therefore the domain portion of that provider's NPPES contact email was FOIA-available.

This resulted in the delivery of 203,939 (177,547 after deduplication) records. Much less than the millions we were hoping for. However, we realized that CMS policy moving forward would begin to require that every hospital and healthcare provider publish a record in the NPPES endpoint file, or eventually, face penalties from CMS. With this court decision as a precedent, replications of this FOIA will result in the contact email domain names for all of the relevant records in NPPES.

This paper reviews the records we received in the FOIA lawsuit and what the data means, and then describes our analysis on interoperability that would have been impossible to conduct without this new data. Specifically, the FOIA'd data allows us to link accountability for contents of websites (in this case, HIPAA medical record request forms) and data that is coded by NPI, (in this case, EHR incentive payments). This allows us to estimate CMS's performance of their statutory duty to promote online transfer of healthcare records to patients.

Background on the endpoint file

During the Obama administration a great policy undertaking resulted in the creation of the "endpoint file" which is an additional data file in the NPPES download that initially contained Direct Addresses, CONNECT IHE (this is an old interoperability protocol that has mostly been retired in favor of FHIR), and email addresses. Recent changes to the NPPES database have added FHIR URLs and removed support for email endpoints. To learn more about what types of endpoints are currently supported, read the data documentation available inside the monthly NPPES data download.

Great effort was undertaken by the original architects of the Direct Protocol to ensure that it was not possible to send SPAM via spoofing using the Direct Protocol. Because it is safe to share Direct Addresses broadly, the provider directories for Direct Addresses were originally intended to be fully public. This would allow patients to find their providers' Direct Addresses easily and then use these to communicate with their physicians and hospitals, instead of using fax machines.

After the Direct Protocol was designed, the Direct Project transmuted into the Direct Trust organization. Direct Trust decided that provider directories for Direct Addressees would be a proprietary asset of the Direct Trust organization. Currently it is not possible to acquire the Direct Trust provider directory without signing a non-disclosure agreement and paying a substantial subscription fee.

It is hard to argue that this is not a form of information blocking. As a result, there has been a long-running policy war between Direct Trust and CMS regarding the open availability of Direct Address endpoints. This culminated in the "Interoperability Final Rule" where it states that providers would be publicly called out for failing to publish their digital endpoints in the NPPES directory. Specifically from page 25584 of the rule:

> *Final Action: After consideration of the comments received, and for the reasons outlined in our response to these comments and in the CMS Interoperability and Patient Access proposed rule, we are finalizing to publicly report the names and NPIs of those providers who do not have digital contact information included in the NPPES system beginning in the second half of 2020 as proposed. Additionally, we will engage in continued public education efforts to ensure providers are aware of the benefits of including digital contact information in NPPES, including FHIR API endpoints, and when and where this information will be posted.*

Eventually, this "name and shame" approach will likely be supplanted with good old fashioned monetary penalties. Normally, CMS begins with "carrot" incentives and then changes to "stick" but so much money has already been spent on funding EHRs and interoperability, that this particular issue will likely draw financial penalties sooner.

This context is critical to understand for this FOIA request. Right now, the limitation on what data we were able to receive was based on the limited number of providers who currently have records in the endpoint file. However, future FOIA requests of the same data will reveal more and more data, making this dataset more valuable over time.

To illustrate this point, the endpoint file from January of 2021 had about 170 thousand rows of data. The endpoint file from August 2021 has 290 thousand rows of data. This increase is due to ongoing efforts by CMS to ensure that all relevant providers have data in the NPPES endpoints file.

The endpoint domain names and the contact email domain names are not the same.

Based on statements that the CMS FOIA office made in our lawsuit, the office does not understand the relationship between the domain names that are available in the endpoint file vs those that are part of contact email addresses in NPPES.

In many cases, these are the same thing. But in most cases they are not.

Only 38,122 records have a shared domain name in the endpoint field as in the contact email field. Most of the time, this is the result of a provider using the same second-level domain name as they use for their email and web addresses, in their endpoint address. Examples of this include records like the following:

Endpoint Domain	**Contact Email Domain (FOIA)**	**Provider Count**
direct.berkeleyeye.com	berkeleyeye.com	43

HISP Services prevent using the endpoint file alone for calculating affiliations

Organizations called Health Information Service Providers, or HISP for short, provide medical practices and hospitals with deployments of the Direct Protocol. Direct addresses look like email addresses but they are not, as they have a specific encryption protocol that ensures that communication across the network is secure from eavesdropping. This is different from email, which can be read by any "transferring computer" that routes the email traffic from one end of the Internet to the other.

For instance, there are 678 different providers with an address of the form:

user_name@directaddress.net

But these providers are not affiliated with one another. They are simply hiring the same HISP, in this case, the domain name directaddress.net is owned by Secure Exchange Solutions.

Using the FOIA'd domain data, we can see that endpoints that use directaddress.net map to 103 distinct healthcare providers. Here are the top 5 rows of those data points:

Endpoint Domain	**Contact Email Domain (FOIA)**	**Provider Count**

directaddress.net	thedermcenters.com	58
directaddress.net	wederm.com	37
directaddress.net	swnamd.com	27
directaddress.net	dermatologyconsultants.org	25
directaddress.net	digestivespecialists.com	23

The directaddress.net domain is one example of a HISP preventing the use of Direct endpoints for the purposes of affiliation analysis.

Other endpoint domain names show similar results. The most notable of these is the 3,499 providers whose affiliations are obscured by *athenahealth.com Direct Addresses. However, athenahealth.com Direct Addresses do allow for some clustering of provider affiliations once they are combined with FOIA data, because the full domain names for athenahealth.com Direct Addresses include a customer identifier number. This gives rise to the following pattern:

Endpoint Domain	**Contact Email Domain (FOIA)**	**Provider Count**
1498.direct.athenahealth.com	harbinclinic.com	262
17389.direct.athenahealth.com	holzer.org	176
5857.direct.athenahealth.com	martinspoint.org	83
1576.direct.athenahealth.com	northwestmedicalcenter.com	74
2340.direct.athenahealth.com	porterhealth.com	69

Even with this small sample size, it is possible to resolve a substantial number of the athenahealth endpoints to specific practices. As the percent of the data that is available under this FOIA grows over time, it will become easier and easier to understand the relationships between providers, organizations, and websites. The task of estimating provider affiliations is a critical first step in many epidemiological studies. Because multiple HISPs are now unblinded, it is now possible to conduct these studies by using the FOIA data, and the usefulness of this approach will only increase with time.

Results

CMS interoperability incentive payments do not yet result in universal access to online records for patients, even among providers who are advertising digital communications to other healthcare providers. But progress has been made, and many of the medical record request forms mention some type of provider-to-patient digital communication.

- We scored the forms associated with 165 domain names. There were 51,780 distinct NPI records connected to these domains.
- We estimated that these providers treat at least 50 million patients. This assigns 1,000 patients per provider (more below).
- 14,068 of these NPIs were paid directly under the interoperability incentive programs.
- The direct payments added up to $691,859,588.34, or 691 million dollars of tax-payer money paid to this cohort of providers to achieve digital interoperability.
- Nearly all of the providers, 51,537, support traditional snail mail access or fax access to patient records.
- 9,413 providers, more than 18 percent of our sample, did not provide an option for online medical record delivery on their request form that would qualify as “interoperable”.
- Those 9,413 clinicians received $149,692,844.29 of interoperability payments.
- We estimate that at least 9.4 million patients (out of 50 million total) are impacted by providers who do not provide online access via their record request forms.

This is a “floor” analysis, we are intentionally choosing values where we can identify the dollar amount of government spending that was definitely wasted. We take this approach in order to state that there is no financial conjecture in this analysis. There are several reasons that the actual number of dollars wasted and the number of patients impacted are higher than we are estimating.

- This study was done on a sample of a sample. First, the FOIA data only includes those providers who published their endpoint data, which is the first of several selection biases. We also did not include all of the released domains in the analysis, instead generally focusing on the domains that have the most providers associated with them. This means that EHR vendors who are especially interoperability friendly (e.g. athenahealth) and their customers are well-represented in this dataset. We can safely assume that healthcare providers that are choosing not to publish their interoperability endpoints in NPPES are likely worse at sharing data with patients too. We also excluded all providers whose forms we could not find on their websites. We believe that most of these biases serve to make those with “good” forms more numerous in our data, than would be found if we could measure every provider in NPPES.
- We are intentionally underestimating patient panel size in this analysis to avoid double counting. The actual impacted patient population could easily be more than double what we are estimating. We chose 1,000 patients per provider because this is below industry standard patient panel estimates.

- Most of the money spent on interoperability incentives were paid to hospitals, who took payments "for" the clinicians who worked at a given hospital. That means that when surgeons, radiologists, anesthesiologists, hospitalists and other doctors who work only in hospitals publish their endpoints, it appears that they have received no EHR incentive payments. Our analysis will capture these EHR incentive payments only when the hospital itself publishes an endpoint. This causes another underestimation of the dollars wasted.
- We are not considering payments made under the Medicaid incentive system, only the Medicare system.

What you can do as a healthcare provider or EHR vendor

You can ensure that your medical record release form has all potential record release channels listed. HHS guidance, which is how HHS provides clarity on the implications of statute, clearly states that if a patient asks for their records to be emailed, they must be emailed. But most patients will assume that the only options they have to access their data are those that are listed on the medical record release form. Only a small percentage of the patients who would prefer to receive their records over email know that they have the right to ask for and receive this.

If you are an EHR vendor, it is time to let your clients know that their medical record release form will be the subject of scrutiny moving forward, and when they fail to give patients proper options of online access, this will eventually become something that impacts the reputation of their organization.

What you can do as a patient

You can review your local hospital or clinic's medical record release form. Ask that they list all record release methods and formats that they support on the form, including email.

What you can do as a citizen

You can call your local representative and ask why CMS paid providers to provide access to online records, while allowing those same providers to represent to patients that only fax and mail are available as methods of access.

You can track and follow regulations on regulations.gov. This link will allow you to see recent regulations, still open for comment that mention HIPAA, the law that is usually under consideration. CMS and HHS regularly change their policies based on regulatory responses from the public.

You can choose to vote for politicians who have extensive experience in healthcare. Wikipedia holds that 24 members of a recent Congress had clinical degrees, including 16 that were medical doctors.

What you can do as a researcher

We've made this FOIA dataset available to the public, and have also published our source code and links to other data required to replicate our analysis. We expect this FOIA dataset to grow as more providers publish their endpoints in NPPES. If you are a researcher, please consider FOIA-ing this data yourself. This can invoke the provisions of eFOIA that ensure that popular datasets become public use files, automatically updated and available on the CMS website.

You can replicate our study, or you can create your own. For inspiration, you can review the many areas related to barriers for patient access to their medical records that we encountered during our study, listed below. Or, this data can be used for other affiliation analyses.

Replicating this analysis requires the use of the following datasets:

- NPPES and the NPPES endpoint file
- EHR Incentive Payments Public Use File

To facilitate more convenient analysis we have published the source code we used to prepare and analyze the data here:
https://github.com/CareSet/NPI_ContactDomain_FOIA_Data

Areas for further study

- If the provider charges for copies/access, and whether or not they list the charges
- If the provider required ID or other types of identity verification
- If spanish or other language versions of the form were available
- If certain types of records are restricted on the form
- If a patient portal was missing on the provider's site
- If patients can fill out and/or send the form electronically
- If patients have to call to request records
- If the provider asks for the purpose of disclosure, and what options are listed
- If the provider restricts certain delivery methods to patients vs providers
- If different forms are required for release to patient vs a provider

Methodology

How we scored HIPAA forms, etc.

- We are only considering one portion of the patient user experience, this medical record request form. Other UX may provide more options, including:
    - Some portals may provide for organization to organization data transfer.
    - Websites may give more options outside of the form.
    - Contacting clinics by phone may result in more options.
    - Organizations may advertise an Android or Apple app, with clinical data integration that has data transfer options.
- Our initial analysis looks at a subset of the released FOIA data. We began by looking at the domains connected to athenahealth endpoints. We then broadened our selection to those domains not associated with athenahealth endpoints. Then we restricted our selection by just looking at domains with certain levels of npi_counts. So, we began narrow, broadened, and narrowed again. This was due to a limitation in resources for this first project. The result is that we've captured domains associated with athenahealth endpoints, plus all of those with npi_count of over 50, about half of those with npi_count 20 and above, and a small percentage of domains with npi_count of less than 20.
- Forms were gathered by doing a site search on the domain using terms like "records request", "medical release form", and/or by browsing the website.
- Many times there is a different web domain than email domain for an organization. This is typically the case when the domain name on the form url is different from the domain name listed in the original dataset. When there is an auto-forward from the email domain to the web domain (example), or search results show that the email domain is listed as contact information in the web domain (example: hhcs.org leading to hamiltonhealth.com), we linked the email domain to the web domain.
- Some organizations have an online form wizard rather than a downloadable form. Because you had to provide valid patient information to begin the wizard process, we excluded these from the analysis.
- Multisites, domains for which we found different forms, were excluded. Multisites also include large websites where many urls (typically websites for different practice locations) would have to be searched. Example.
    - One exception was if a domain included multiple locations with different urls, but it appeared that there was one main request form for the entire domain. Example.
    - Another exception was if a site listed a small number of multiple forms, but we found that all the forms were identical, or varied only slightly, such as having a different site name/address at the top. Example.
- Some .edu domains had medical request forms for their student health plans that were distinct from their public medical request forms, as a result we excluded most .edu sites. Example. We did include some very large academic medical centers, like mayo.edu where this was not an issue we encountered.
- Very often, we could not find a form for a particular domain. Sometimes, the domain did not lead us to a valid working url. Sometimes the provider website did not have a form.

- If a mailing address or fax number was included in the "release to" section of a form, we included these as options. These were included even if they were not mentioned in a "delivery method" section. Fax example.
- If a form did not specify any type of method for sending the information, we marked all options as "0". Example.
- If certain delivery methods were only available for particular types of records, for example CDs only for imaging records, these were not counted. Example.
- It is possible that a given provider might have an outdated email address in their NPPES record. However, this is unlikely to be common for two reasons. First, a provider is legally required, under HIPAA, to keep their NPPES information current. Second, when a provider changes organizations they will lose access to their previous employer's email, and then when they go to edit their records in NPPES/PECOS, they will need to update their email address to login. All providers who bill Medicare would be required to do this to direct their Medicare payments to their new employers, ensuring that for the vast majority of healthcare providers, this email is likely to provide the correct affiliation information. There is no way to validate this hypothesis (that the affiliations are valid over time), however, without ongoing access to this dataset, and comparing it with other changes in the NPPES record. Our spot checking against provider registries indicate that the affiliations are currently valid.
- It is possible that we are scoring individuals or organizations that are not clinicians. However, interoperability payments only went to clinicians. Given that only clinicians were eligible for these payments, the financial analysis is still accurate.

Relevant Previous Work

As far as we know we are the first to link HIPAA medical record request forms to EHR incentive payments for analysis.

But we feel that Ciitizen's work scoring whether providers respond reasonably to medical records request serve as both precedent and inspiration. Further there is a long-standing vocal movement in the patient community to ensure data access, which is embodied by Dave deBronkart's (e-patient Dave) continued cry of "Gimme my damn data". This movement is a continued reminder of how much this matters to anyone who is sick and the public at large.

Our study approach shared design components employed by the Office of the National Coordinator for Health Information Technology (ONC) in this study. Because our study was powered by the FOIA data, we were able to triple the number of medical release forms examined, from the number evaluated in the ONC study.

- Lye CT, Forman HP, Gao R, Daniel JG, Hsiao AL, Mann MK, deBronkart D, Campos HO, Krumholz HM. Assessment of US Hospital Compliance With Regulations for Patients' Requests for Medical Records. JAMA Netw Open. 2018 Oct 5;1(6):e183014. doi:

10.1001/jamanetworkopen.2018.3014. Erratum in: JAMA Netw Open. 2018 Dec 7;1(8):e186463. PMID: 30646219; PMCID: PMC6324595.

- deBronkart D, Eysenbach G Gimme My Damn Data (and Let Patients Help!): The #GimmeMyDamnData Manifesto J Med Internet Res 2019;21(11):e17045 URL: https://www.jmir.org/2019/11/e17045 DOI: 10.2196/17045
- Health Care Provider Compliance with the HIPAA Right of Individual Access: a Scorecard and Survey (Revised) Deven McGraw, Nasha Fitter, Lisa Belliveau Taylor medRxiv 19004291; doi: https://doi.org/10.1101/19004291
- Improving the Health Records Request Process for Patients Insights from User Experience Research. (n.d.). https://www.healthit.gov/sites/default/files/onc_records-request-research-report_2017-06-01.pdf

Updates:

- Initial analysis released on Aug 17th 2021
- Added background on the endpoint file Aug 17th 2021
- Adding information blocking, and channel of access comments, and references Aug 19th, 2021
- Adding validation information and new references Aug 23rd, 2021