Oregon All Payer All Claims (APAC) Program

Application for Limited Data Files

APAC-3

This application is used to request limited data sets. If you would like to discuss APAC data in relation to your project prior to submitting this application, please contact apac.admin@state.or.us with a brief description of the project and your contact information. OHA will have someone contact you to help determine if APAC is appropriate for your project and, if so, which data elements may be needed.

PROJECT INFORMATION

Project Title: Pharmacy Deserts and Cross-Subsidization

Principal Investigator: Ranie Lin

Title of Principal Investigator: PhD Student

Organization: Princeton University

Address: Julis Romo Rabinowitz Building, Washington Rd

City: Princeton State: NJ Zip Code: 08544

Telephone: 4692376981

Email: ranie.lin@gmail.com

SECTION 1: PROJECT SUMMARY

1.1 Project Purpose: Briefly describe the purpose of the project. You may submit a separate document that details the project's background, methodology and analytic plan in support of your request for APAC data elements.

The purpose of our project is to analyze spatial inequality in access to pharmacies and understand the economic mechanisms behind pharmacies' operating decisions. Please see the attached proposal for our detailed research plan.

1.2 Research Questions: What are the project's key research questions or hypotheses? If this project is research and has been approved by an Institutional Review Board (IRB), the research questions must align with the IRB approval documentation. If needed, a more detailed response may be submitted as a separate file.

- Note: APAC staff will use your response to this question to determine the minimum data elements necessary for this project, in accordance with the HIPAA minimum necessary standard. The research questions should be specific enough to justify the need for each data element beyond identifying it as a "potential confounding variable."

Our project's key research questions include:
- How does access to pharmacies differ across low and high income communities and rural versus urban areas?
- How much are pharmacies reimbursed for various types of prescription drugs?
- Are pharmacies' front-end retail products an important component of pharmacy profits, possibly cross-subsidizing low margins from filling prescription drugs?

Please see the attached proposal for our detailed research plan.

1.3 Products or Reports: Describe the intended product or report that will be derived from the requested data and how this product will be used. If needed, a more detailed response may be submitted as a separate document with this application.

The intended product is an academic research paper that will be submitted for publication in a peer-reviewed economics journal.

1.4 Project Timeline: What is the timeline for the project?

Anticipated Start Date: 05/15/2024

Anticipated Publication/Product Release Date: 01/01/2027

Anticipated End Date: 01/01/2027

1.5 Data files may not be released or reused beyond the terms of the data use agreement resulting from this application regardless of funding source or other obligations of the principal investigator, organization or research team.

☒ I understand this limitation and agree that data files or work products will not be shared at less than an aggregated, de-identified level.

☐ I understand this limitation and request approval to share data files or work products at a potentially re-identifiable level as follows:

SECTION 2: PROJECT STAFF

2.1 Project Staff: Please list all individuals in addition to the principal investigator who will have direct or indirect access to the data. This must include any contractors or other third parties with access to the data.

Name: Renjie Bao Email: renjie.bao@princeton.edu	Project role: Collaborator
Name: Email:	Project role:
Name: Email:	Project role:
Name: Email:	Project role:
Name: Email:	Project role:
Name: Email:	Project role:
Name: Email:	Project role:

Attach additional sheets as needed.

2.2 Technical Staff: Please list any additional staff who will be maintaining the data file(s) or otherwise assisting in the transfer or receipt of the data files. Files will not be transferred to anyone who is not listed on this application as either project staff or technical staff.

Name: Email:	Technical role:
Name: Email:	Technical role:

Attach additional sheets as needed.

SECTION 3: DATA REQUEST

3.1 Purpose of the Data Request:

a. Listed below are the purposes for which OHA may share APAC data. Please choose the category in which your project falls under (***choose only one***).

- [■] Research (refer to 45 CFR 164.501 for definition)
- [ ] Public health activities as defined in 45 CFR 164.512(b) by the state or local public health authority
- [ ] Health care operations as defined in 45 CFR 164.501
  - Covered entity as defined in 45 CFR 160.103? [ ] Yes [ ] No
- [ ] Treatment of patient by health care provider as defined in 45 CFR 164.506 (c)(2)
  - Covered entity? [ ] Yes [ ] No
- [ ] Payment activities performed by covered entity or health care provider as defined in 45 CFR 164.506 (c)(3)
  - Covered entity? [ ] Yes [ ] No
- [ ] Work done on OHA's behalf by a Business Associate as defined in 45 CFR 160.103

b. Describe how the project falls into the category chosen above.

We are PhD students who will be using the APAC data to write an academic research paper.

3.2 Direct identifiers. What level of data identifiers are you requesting (***choose only one***)? Reference the Data Elements Workbook for the categorization of data elements.

- [ ] De-identified (as outlined in 45 CFR 164.514(e)) protected health information
- [■] Limited, potentially re-identifiable data elements
- [ ] Restricted direct identifiers (member name, address, date of birth, etc.) *Please note:* Direct identifiers are only released under special circumstances that comply with HIPAA requirements, and will require specific approvals, such as IRB approval, patient consent and/or review by the Oregon Department of Justice.

3.3 Human Subjects Research: IRB protocol and approval are required for most research requests for limited data elements. Not obtaining IRB approval or waiver in advance may delay approval of the data request. **The research questions reported in 1.2 of this application must match the documentation supporting the IRB approval received or the IRB approval will not be accepted for this data application.**

The IRB application should indicate that APAC data contains sensitive personal health information and is subject to HIPAA regulations.

a. Does the project have IRB approval for human subjects research or a finding that approval is not required?

☐ Yes ■ No

If no, briefly explain why you believe that this project does not require IRB review.

Our project does not involve the use of direct individual identifiers and is not human subjects research. We are attaching the IRB waiver to our application.

If an IRB reviewed the project, include the IRB application and approval/finding memo with the submission of this APAC-3 and complete parts b-e below.

☐ IRB application and approval memo are attached.

b. Describe how this application is within the authority of the approving IRB.

c. Describe why the project could not be practicably conducted without a waiver of individual authorization (a waiver of individual authorization is provided by the IRB in cases in which the researcher does not need written authorization from participants to use their PHI):

d. On what date does the IRB approval expire?

SECTION 4: DATA ELEMENTS

4.1 Narrowing Data Needs: Refer to the APAC Data Dictionary for detailed information about the data elements. In compliance with HIPAA regulations, you will only receive data elements that are adequately justified. This means APAC will only provide the minimum necessary data required for the project as represented in the research questions, protocol and IRB approval.

a. What years of data are requested? 2011 through 2021 are currently available.

2019, 2020, 2021, 2022

b. What payer types are requested? Check all that apply

☑ Commercial ☑ Medicaid ☑ Medicare Advantage

c. What types of medical claims are requested? ☐ All

☐ Inpatient hospital ☐ Emergency department ☐ Outpatient
☐ Ambulatory surgery ☐ Ambulance ☐ Transportation
☐ Hospice ☐ Skilled Nursing Facility ☐ Professional

d. Demographic data limitations

1. Gender ■ All ☐ Male ☐ Female

2. Age ■ All ☐ Only 65+ ☐ Only 18 and younger ☐ Other (Specify age range)

e. Will data requested be limited by diagnoses, procedures or type of pharmaceutical? Add additional sheet if needed.

Diagnoses, indicate ICD 9 and ICD10 codes to include:

Procedures, indicate CPT to include:

Pharmaceuticals, indicate NDC or therapeutic classes to include:

f. APAC has a small number of out-of-state residents included, most often through PEBB or OEBB coverage. Do you want to include out-of-state residents? ■ Yes ☐ No

4.2 Data Element Workbook: Complete the Data Element Workbook to identify specific data requested.

Data Element Workbook completed and attached, including justifications for each element requested. ☑

SECTION 5: DATA MANAGEMENT & SECURITY

5.1 Data Reporting: APAC data or findings may not be disclosed in a way that can be used to re-identify an individual. Data with small numbers – defined as values of 30 or less (n≤30) or subpopulations of 50 or fewer individuals (n≤50) – cannot be displayed in findings or outputs derived from APAC data. Please describe the techniques you will use to prevent re-identification when findings or outputs result in small numbers or subgroups (e.g. aggregation, cell suppression, generalization, or perturbation).

In cases of data with small numbers, we will only present results after those figures have been aggregated to a larger unit or geography to avoid the possibility of re-identification.

5.2 Data Linkage: OHA seeks to ensure that APAC data cannot be re-identified if it is linked or combined with data from other sources at the record, individual or address level. Requesters are strongly encouraged to consult with APAC staff regarding linking APAC data with other data prior to submitting a data request. Health Analytics prefers to conduct APAC data linking in-house and share only encrypted identifiers with data requesters.

a. Does this project require linking to another data source?

☐ Yes ■ No

If yes, please complete parts b-d below.

b. At what level will data be linked?

☐ Address ☐ Facility ☐ Individual person/member

☐ Individual provider

c. If required to link

☐ Authorized to provide data for linking at OHA

☐ Not authorized to provide data for linking at OHA

☐ Unknown

d. Describe and justify all necessary linkages, including the key fields in each data set, how they will be linked, the software proposed to perform the linkage and why it is necessary.

e. Describe in detail the steps will you take to prevent re-identification of linked data.

5.3 Data Security (required for all applications):

a. Attach a detailed description of your plans to manage security of the APAC data including:
   - Designation of a single individual as the custodian of APAC data, either the principal investigator or staff listed in Section 2 of this application, who is responsible for oversight of APAC data, including reporting any breaches to OHA and ensuring the data are properly destroyed upon project completion.
   - A security risk management plan applicable to APAC data that includes:
     - Secure storage in any and all mediums (e.g., electronic or hard copy)
     - Procedures to restrict APAC data access to only those individuals listed on the data use agreement
     - User account controls, i.e., password protections, maximum failed login attempts, lockout periods after idle time, user audit logs, etc.
     - Confirmation of training for personnel on how to properly manage protected health information in all formats
     - Protection of derivatives of APAC data at the re-identifiable level
     - If applicable, procedures for handling direct identifiers, such as allowing access on a 'need to know' basis only and minimizing risk by storing identifiers separately from other APAC data
     - Procedures for identifying, reporting and remedying any data breach
   - Statement of compliance with HIPAA and the HITECH Act
   - Electronic device protections, i.e., anti-virus or anti-malware software, firewalls, and network encryption

b. Record level or derivative data that can be re-identified must be destroyed within 30 days of the end of the data use agreement, in a manner that renders it unusable, unreadable or indecipherable. What are your plans for destruction of the dataset and any potentially identifiable elements of the data once the data use agreement has expired?

All record-level or derivative data that can be re-identified will be permanently deleted from Citadel, Princeton's secure research infrastructure environment that will be used to store the data.

SECTION 6: COST OF DATA

Because each data set is unique, cost can be determined only after the specific data elements are finalized. APAC staff will then review your request and estimate the number of hours required to produce and validate the data. APAC is currently requiring reimbursement for the cost of file transfer only ($890 per request). Payment must be received before the data will be provided. APAC staff will provide an invoice to facilitate payment. OHA's W-9 is available on request.

SECTION 7: CHECKLIST AND SIGNATURE

7.1 Checklist: Please indicate that the following are completed:

- ☒ I acknowledge that payment will not be refunded if OHA fulfills the data request, but the receiving entity does not have the capability to import or analyze the data
- ☒ All questions are answered completely
- ☒ Data Element Workbook is attached to email or printed application
- ☒ IRB application with approval/finding memo is attached to email or printed application, if applicable
- ☒ Data privacy and security policies for the requesting organization, and any third-party organizations, are attached to the email or printed application

7.2 Optional Racial Justice Addendum: Please see the last two pages of this form for options if data will be used to eliminate racial injustice.

- ☐ I am interested in this option
- ☒ This option does not apply to my data request

7.3 Signature: The individual signing below has the authority to complete this application and sign on behalf of the organization identified in Section 1. By signing below, the individual attests that all information contained within this data Request Application is true and correct.

Signature Date 05/15/2024

Printed name Ranie Lin

Title PhD Student

Return the completed form with required attachments to APAC.Admin@odhsoha.oregon.gov.

Optional APAC Addendum: Using APAC Data to Eliminate Racial Injustice

Requestors may complete this optional section if their project will identify concrete actions to eliminate health inequities stemming from historical and contemporary injustices and the inequitable distribution of resources and power (see Health Equity definition on next page). For projects that inform such solutions, and **do not simply document disparities**, the Director of the **Office of Health Analytics** may, at their discretion, offer one or more of the following incentives:

- Priority processing of requestor's application
- Waiver of fees
- Priority production of data files
- Technical assistance from APAC analysts
- Access to enhanced race and ethnicity data in the future. (Race/ethnicity data in APAC are currently limited because entities that submit administrative data to APAC do not generally include race/ethnicity information.)
- Other provisions that the Director of Health Analytics may find appropriate

Receipt of any of these incentives requires requesters to deliver to the Office of Health Analytics a document fully describing the analytic methods at the conclusion of the relevant analyses, including:

- Commercial off-the-shelf applications used
- Grouping and aggregation methods
- Algorithms and calculations
- Use of code sets that are proprietary to a third party not associated with the project
- Copies of programming code attached in an appendix

The Office of Health Analytics will compile a compendium of analytic methods and make this freely available on the APAC web site. Requestors are also encouraged to submit copies of publications or products using the APAC data for posting on the APAC web site. See below for additional information and application instructions.

Using APAC Data to Eliminate Health Inequities

Problem: Health inequities due to institutional racism and racial injustice

Solution: Develop methods for using APAC data to eliminate institutional racism and racial injustice.

Goal: Eliminate institutional racism and racial injustice, including discrimination based on the intersections of race, ethnicity, language and disability.

Rationale: OHA recognizes that historical and contemporary racial injustice is a root cause of health inequity. APAC and its users, who have subject matter expertise, infrastructure, and staffing sufficient to use the large and complex data files, comprise a community of privilege. As such, APAC has an obligation to use its privilege to confront institutional racism and racial injustice, within OHA specifically and across Oregon. The APAC community has a tremendous wealth of research expertise that could develop novel methods for using APAC data to document racial injustice and identify opportunities to eliminate it.

Instructions: In a separate attachment, describe in detail:

- How requestor's research will help requestor's organization and OHA document racial injustice **and** identify opportunities to eliminate it. Requestor's description must be thorough and as specific as possible and should describe how the research findings will be consistent with OHA's efforts to achieve true Health Equity (see definition, below). **Simply documenting disparities is not sufficient**.
- How requestor's research will be explicitly clear and open about the methods used, widely replicable, and not proprietary to requestor's organization or to a third party. Note that this does not preclude requestor's use of necessary codes sets, such as CPT codes, that are proprietary to a third party and available for license.
- How requestor's organization will freely share the key findings.

A note on intersectional research into inequities based on race, ethnicity, language and disability: Researchers are encouraged to consider an intersectional approach that encompasses language and disability when researching strategies to eliminate racism and racial injustice. However, administrative claims data submitted to APAC generally do not include data on language or disability. APAC includes some race and ethnicity data, but it encompasses less than half of the people in the database. To mitigate these limitations, OHA staff may be able to provide assistance to selected applicants interested in intersectional approaches, as staff resources permit.

Health Equity Definition

Oregon will have established a health system that creates health equity when all people can reach their full health potential and well-being and are not disadvantaged by their race, ethnicity, language, disability, gender, gender identity, sexual orientation, social class, intersections among these communities or identities, or other socially determined circumstances.

Achieving health equity requires the ongoing collaboration of all regions and sectors of the state, including tribal governments to address:

- The equitable distribution or redistributing of resources and power; and
- Recognizing, reconciling and rectifying historical and contemporary injustices.

Pharmacy Deserts, Cross-Subsidization, and Spatial Inequality

Renjie Bao[†] Ranie Lin[‡]

May 17, 2024

Abstract

This is a proposal for our project on pharmacy deserts and cross-subsidization. We discuss our research questions, intended methodological approach, and how we plan to use the Oregon APAC data.

[†]renjie.bao@princeton.edu
[‡]ranielin@princeton.edu

1. Introduction

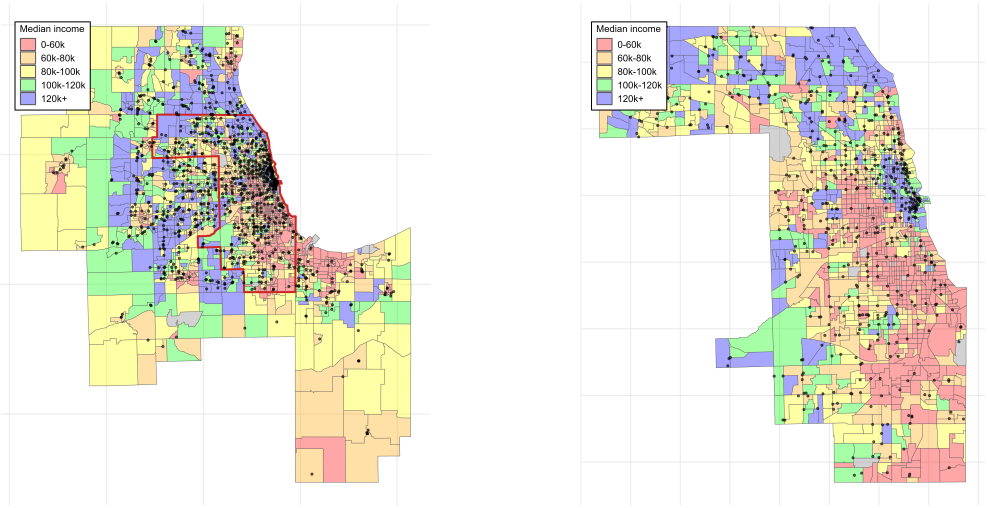
Pharmacies dispense prescription drugs, supply emergency doses of life-saving medications, and provide vaccinations and contraceptives. In recent years, there has been growing concern regarding equity of access to these crucial health services. Both pharmacy chains and independent pharmacies, citing low profits, are closing stores in low-income, minority neighborhoods, creating what public health experts refer to as a "pharmacy desert" crisis (Qato et al. 2014). As a motivating example, Figure 1 shows that the concentration of pharmacies across the Chicago metropolitan area substantially decreases toward the lower-income South Side neighborhoods.

We begin by building robust empirical facts related to pharmacy access and inequality. Controlling for population density, a proxy for rural or urban areas, we find that a $10k decrease in median household income is associated with an average increase in the distance to the nearest pharmacy by 0.2km. Inequality in access is particularly strong in rural areas with population densities in the bottom four deciles, which comprise 37.7% of the total population and 98.4% of the total land area. For example, in the least dense areas, a $10k decrease in median household income is associated with an increase in the distance to the nearest pharmacy by more than 0.8km. We are also working to build more facts related to recent concerns on pharmacy closures and their impacts on health outcomes, such as prescription adherence (Gannaway 2016).

From a *normative* perspective, we argue that there is a fundamental inefficiency generated by pharmacies' entry and location choices due to the externality associated with consumers' private travel costs. Entry into an area with few pharmacies could reduce the travel time to pharmacies for local residents, but this social benefit may not be fully internalized by pharmacies for two reasons. First, pharmacies are not able to freely set prices for prescription drugs but instead must bargain with private insurers over reimbursement rates (or take as given rates set by the government, in the case of public insurance). An imperfect price mechanism could thus result in places where consumers have a high willingness to pay for medication but low access to pharmacies. Second, some consumers may have acute health symptoms, resulting in an urgent need for medication and an especially large private cost of travel that pharmacies cannot profit from. Both factors could amplify the inefficiency in pharmacy entry.

We examine the mechanisms behind pharmacy deserts from an important but overlooked dimension: cross-subsidization between the market for prescription drugs, dispensed in pharmacies, and the market for retail products (e.g., wellness and beauty, food and beverage, etc), referred to as the "front end" of the pharmacy. Pharmacies report that front-end margins are often substantially larger than prescription drug margins, which could potentially offset inefficiencies by providing an additional incentive for pharmacy entry. A natural spillover occurs via foot traffic when consumers who arrive at pharmacies to fill prescriptions also find it appealing to purchase front-end products conveniently located in the same store, strengthening the

Figure 1: Pharmacy deserts in the Chicago metropolitan area



(a) Chicago-Naperville-Elgin MSA (b) Cook County, IL

Notes: The maps show the distribution of chain pharmacies across the Chicago MSA in panel (a) and across Cook County in panel (b). Geographic units are Census tracts colored by median household income. Each black dot corresponds to a chain pharmacy. The lower-income South Side has less access to pharmacies than higher-income neighborhoods.

cross-subsidization effect. However, in low-income communities, the ability of pharmacies to extract high margins from front-end sales may be diminished, exacerbating spatial inequality in pharmacy access.

To quantify the cross-subsidization mechanism and its effect on pharmacy access, we build and estimate a spatial model of demand for pharmacies. In our model, consumers have distaste for travel but can purchase both front-end goods and prescription drugs in the same retail pharmacy without incurring additional travel costs, leading to cross-market spillovers. We plan to estimate our model using data from Oregon, a state where pharmacy deserts have become a major concern, especially in rural areas.[1] Using data from the Oregon All-Payer All-Claim Reporting Program (APAC) and Safegraph foot traffic and retail spending data, we can construct pharmacy-level market shares for both medication and retail sectors. We plan to conduct two main counterfactual analyses with our estimated model. First, we will examine the effects of increasing reimbursement rates for prescription drugs (e.g., Medicaid rates), which is the policy

[1] For example, see https://www.opb.org/article/2024/02/16/rural-pharmacies-continue-to-close-as-legislature-takes-another-look-at-regulation/ for a discussion of rural pharmacy closures.

proposal advocated by the National Community Pharmacists Association (NCPA). Second, to highlight the importance of cross-subsidization, we plan to evaluate the effects of pharmacy-retail mergers, such as the recent acquisition of Target's pharmacy business by CVS Health.

Literature. Our study of pharmacy access aims to formalize the economic mechanisms underpinning pharmacy deserts and inequality, a pattern that has been highlighted in the public health literature. In a case study of Chicago, Qato et al. (2014) find that pharmacy deserts are disproportionately located in segregated, Black communities, low-income communities, and federally designated Medically Underserved Areas. Examining direct health outcomes, Gannaway (2016) shows that access to pharmacies affects medication adherence; in particular, pharmacy openings cause an increase in local patients' measures of adherence by roughly 2%. A smaller literature examines the effect of pharmacy competition and market structure on regional access (Kim 2023). More broadly, this phenomenon relates to the discussion of regional inequality in access to local amenities such as high-quality food (Allcott et al. 2019) and hospitals (Lindrooth et al. 2018; Moghtaderi et al. 2020).

We demonstrate how the distribution of pharmacies and inequality in access is affected by cross-subsidization, a theory that has long been studied since Faulhaber (1975). Several papers have analyzed the mechanism of foot-traffic spillovers. Wu (2024) provide empirical evidence on the strategic value for retail pharmacies to grow consumer traffic. In a similar spirit, Shoag and Veuger (2018) document that larger retailers produce significant positive spillovers on other local businesses using evidence from big-box bankruptcies. Thomassen et al. (2017) develops and estimates a demand model incorporating cross-market complementarities in the context of supermarket pricing. Our project contributes to this literature by analyzing the interaction between cross-subsidization and pharmacy deserts, a new and policy-relevant setting with important equity implications.

2. Data description

Our analyses will utilize several datasets that cover both pharmacies' prescription drug and front-end sectors. Together, these data provide comprehensive information on pharmacy locations, foot traffic and spending, retail prices, prescription drug claims, and demographics. We provide a brief description of each data source.

SafeGraph spend and foot traffic. The SafeGraph spend and Advan foot traffic datasets, accessed through Dewey, provide hourly foot traffic and daily credit card spending data for a wide range of locations in the U.S. from January 2019 onwards, including more than three million retail stores. We also observe counts of visitors' home Census block group for each location on a weekly basis, which allows us to generate pharmacy and retail market shares at the granular

blockgroup-week level. For our current analysis, we are using a cross-section of the data at the Census block group level from the last week of September 2023.

National Provider Identification (NPI) database. We obtain supplemental information on pharmacies from the Centers for Medicare & Medicaid Services (CMS) NPI database. We observe each pharmacy's address, activated status, and the date of deactivation (if applicable), which will allow us to analyze the evolution of pharmacy deserts over time.

Oregon All-Payer All-Claim Reporting Program (APAC). The Oregon APAC dataset provides administrative data on pharmacy claims. For each claim, we observe the pharmacy identifier and NPI, prescription fill date, National Drug Code, refill information, payment information (payer payment and out-of-pocket payment), member demographics, and home location up to the member's zipcode. From these data, we are able to construct measures of medication adherence, individual-level pharmacy access, and market shares for pharmacies' medication sectors.

Nielsen Retail Scanner data. We obtain retail product prices in drug stores and other retailers from the Nielsen Retail Scanner Data. For each store-UPC bar code pair, we observe the retail type (drug store, convenience, food store, or mass merchandiser), weekly price, quantity, and location at the county level. We use observations from 2019, the most recent year before the Covid outbreak that is available to us.

American Community Survey (ACS). To complement our analysis, we leverage publicly available data from the Census ACS, which provides block group level demographic information including population size, population density, and median household income.

3. Pharmacy access and spatial inequality

We provide empirical patterns and descriptive evidence of pharmacy deserts, cross-subsidization, and spatial inequality. In subsection 3.1, we examine the correlation between pharmacy access and income, highlighting that pharmacy deserts are especially prominent in low-income, rural areas. In subsection 3.2, we focus on the role of pharmacies' front-end retail sectors and their relatively high profit margins. Finally, we provide evidence of cross-market spillovers from the prescription drug to retail sectors in subsection 3.3.

3.1. Pharmacy deserts

We examine the correlation between income and pharmacy access at the Census block group level with the regression

$$dist_i = \sum_{k=1}^{10} \beta_1^k density_i^k + \sum_{k=1}^{10} \beta_2^k \left(income_i \times density_i^k \right) + \alpha_{s(i)} + \varepsilon_i, \tag{1}$$

where each observation i is a block group. We construct the dependent variable $dist_i$ as the distance from the centroid of each block group to the nearest branded pharmacy.[2] This specification allows us to flexibly control for rural and urban characteristics using a set of dummy variables $density_i^k$, which correspond to ten decile bins of block group population density indexed by $k = 1, \ldots, 10$. The key regressor, $income_i$, is the block group level median household income. We also include state fixed effects $\alpha_{s(i)}$ to account for state-level unobservables.

We are interested in the set of coefficients $\{\beta_1^k, \beta_2^k\}_{k=1}^{10}$, whose estimates are reported in Figure 2. In panel (a), we show that distance to the nearest pharmacy declines with population density. For example, people living in the least dense block groups must travel 20km further to access a pharmacy than those living in median-density block groups. More strikingly, we show substantial inequality in pharmacy access by income—particularly for more rural areas—in panel (b). Conditional on population density, low-income block groups have significantly worse pharmacy access compared to their high-income counterparts; moreover, the degree of inequality increases as areas become less urbanized. For the least dense block groups, a \$10k increase in median household income shortens the distance to the nearest pharmacy by 0.8km. We note that block groups with population densities at the bottom four deciles, where the effect of income on access is significant, comprise 37.7% of the total population and 98.4% of the total land area in the U.S.

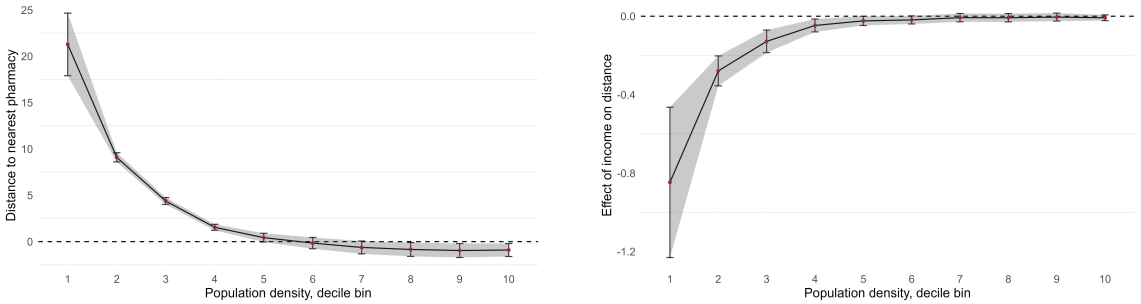
Having established empirical facts about pharmacy deserts, we also hope to study health consequences. A potentially important outcome is medication adherence, as patients with low access to pharmacies may be more inclined to skip prescriptions. Although we do not directly observe prescription skipping, we can impute the level of prescription adherence from APAC claims data, which we plan to later analyze after gaining data access.

3.2. Pharmacy front-end stores

We proceed to study pricing patterns by pharmacies' front-end retail sectors. Specifically, we provide suggestive evidence of relatively high profit margins of retail goods sold by pharmacy front-end stores in Figure 3 using the 2019 Nielsen Retail Scanner data. For nine randomly

[2] Branded pharmacies include both national and regional/local chains. We are working on completing the pharmacy sample by adding independent pharmacies from the NPI database.

Figure 2: Pharmacy access, population density, and income



(a) Effects of density, β_1^k

(b) Effects of income given density, β_2^k

Notes: We plot estimation results for regression specification (1). For each block group population density decile, we report the average distance (net of fixed effects) to the nearest chain pharmacy, β_1^k, in panel (a) and the effect of median household income on distance to the nearest pharmacy, β_2^k, in panel (b). Error bars depict 95% confidence intervals; standard errors are clustered at the state level.

selected product modules, we plot the three quartiles of the distribution of prices relative to drug store prices at the UPC-county level for mass merchandisers, food stores, and convenience stores, respectively. To avoid compositional differences, we filter the sample so that all the products in the corresponding county are sold by all four types of retailers. The results suggest that pharmacies sell most retail goods at substantially higher prices than mass merchandisers and food stores. The retail prices in pharmacies are comparable to convenience stores, which are commonly believed to sell products at high markups. We conclude that profit margins at pharmacy front-end stores are likely to be high and important for their entry decision.

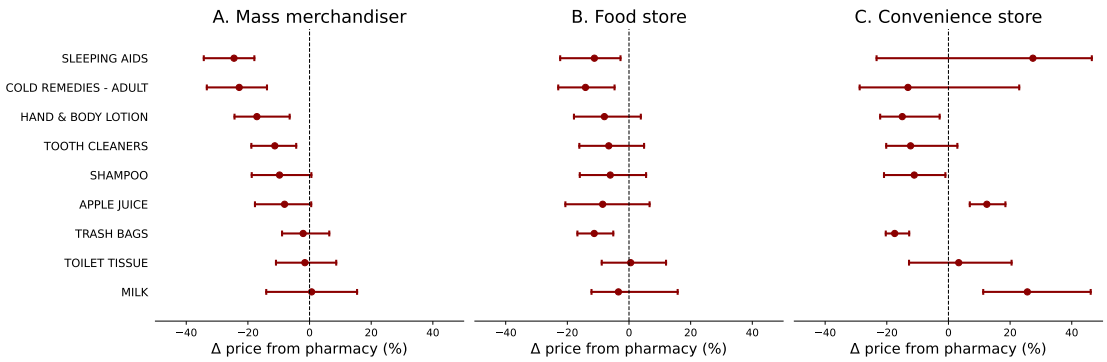
3.3. Cross-market spillovers

Finally, we provide evidence of pharmacy-to-retail spillovers, a core component of our cross-subsidization mechanism. We estimate

$$\log(retail\_transactions_{mt}) = \beta \log(visits_{mt}) + \alpha_m + \delta_t + \varepsilon_{mt}, \tag{2}$$

with observations m, t at the pharmacy store by week level, controlling for pharmacy and time fixed effects through α_m and δ_t. To identify the spillover effect, we instrument for $visits_{mt}$ using the initial roll-out of the Covid-19 vaccine, which we measure as the county by week percentage of the population vaccinated in 2021. The key identifying assumption is that the initial roll-out of the Covid-19 vaccine is correlated with foot traffic to pharmacies but is otherwise unrelated to

Figure 3: Comparison of retail prices between pharmacies and other retailer categories



Notes: We construct the quantity-weighted average price for a large set of products (UPC) at the county × retailer category level in 2019. We then group these products into nine randomly selected product modules from NielsenIQ. In each panel, we report the three quartiles of the distribution of prices relative to drug store prices within each product module for mass merchandisers, food stores, and convenience stores, respectively. We fix the compositional effect by filtering the sample to products that are sold by all four types of retailers for each product-county pair.

demand for retail products after controlling for pharmacy and week fixed effects. We argue that this is a plausible assumption as initial take-up of the vaccine was largely driven by exogenous, regional availability constraints.

We report the estimation results in Table 1, where we execute regression equation (2) with three variations of the instrument. We take regression specification (1), which uses both the first and second Covid-19 doses as instruments, as our baseline. First-stage results suggest that roll-out of the Covid-19 vaccine is strongly associated with pharmacy visits. Our IV results show that a one percent increase in medication-related pharmacy visits increases front-end retail product transactions by 0.2998%. Furthermore, our results are robust to different sets of instruments, where the retail transaction elasticity with respect to pharmacy visits varies between 0.2097 and 0.5352. We therefore conclude that there is a significant, positive within-pharmacy spillover from the medication sector to the retail sector, which is likely to be internalized by pharmacies when making pricing and entry decisions.

4. Model of prescription drug and retail sectors

Motivated by the empirical facts outlined in section 3, we plan to build a spatial demand model for pharmacies that incorporates the core idea of cross-subsidization and cross-market spillover. We then discuss how we plan to identify the model parameters from our data.

Table 1: Pharmacy-to-retail spillover

	Dep var: *log retail transactions*					
	(1)		(2)		(3)	
	IV	First stage	IV	First stage	IV	First stage
log pharmacy visits	0.2998		0.2097		0.5352	
	(0.0494)		(0.0523)		(0.0677)	
Instrument						
first dose	Y	0.0126	Y	0.0156		
		(0.0004)		(0.0004)		
second dose	Y	0.0074			Y	0.0144
		(0.0005)				(0.0004)
Fixed effects						
pharmacy	Y	Y	Y	Y	Y	Y
week	Y	Y	Y	Y	Y	Y
Number of obs.	625,673	629,083	625,673	629,083	625,673	629,083
R2	0.0386	0.9337	0.0613	0.9337	0.0381	0.9336
F statistic	329.70	1029.40	330.26	1827.44	323.26	1122.02

Notes: This table reports estimation results of regression specification (2) using county-level weekly Covid vaccination rates (percentage of population) as an instrument for pharmacy visits. We show the estimation with three variations of the instruments: (1) both first and second doses, (2) first dose only, and (3) second dose only. For each specification, we report the first stage regression alongside IV regression results. Standard errors are reported in parenthesis.

4.1. **Model**

Market structure. We focus on two sectors, retail $\mathcal{R}$ and medication $\mathcal{M}$, and consider two types of firms, pure retailers $r \in \mathcal{R}$ that only sell retail products and pharmacies $m \in \mathcal{M}$ that sell retail products and fill drug prescriptions, which we refer to as medication products. We index stores by $j \in \mathcal{J} := \mathcal{R} \cup \mathcal{M}$ and denote its location as $\ell(j)$. For simplicity, we assume products are representative in both sectors with endogenously chosen retail prices $\{p^R_{jt}\}$ and exogenous retail and medication qualities $\{\xi^R_{jt}, \xi^M_{mt}\}$.[3]

Demand. Consumers $i \in \mathcal{I}$ residing in location $\ell(i)$ have exogenous health status $h_{it} \in \{0, 1\}$ at time t, where $h_{it} = 0$ if the consumer is sick and $h_{it} = 1$ if the consumer is healthy. The mean utility (net of travel costs) that consumers obtain from consuming retail and pharmacy products,

[3] We assume pharmacies cannot choose drug reimbursement rates. Instead, they take the reimbursement rate p^M_{mt} as given. Our model thus abstracts from the complex bargaining process between pharmacies and insurers, which is not the focus of this paper.

respectively, is

$$\delta^R_{ijt} = \alpha_i p^R_{jt} + \underbrace{\mathbf{X}^R_{jt}\boldsymbol{\beta}^R_i}_{\text{obs. characteristic}} + \underbrace{\xi^R_{jt}}_{\text{unobs. quality}} \tag{3}$$

$$\delta^M_{imt} = \left(\underbrace{\mathbf{X}^M_{mt}\boldsymbol{\beta}^M_i}_{\text{obs. characteristic}} + \underbrace{\xi^M_{mt}}_{\text{unobs. quality}} \right) \times \underbrace{\mathbb{1}\{h_{it} = 0\}}_{\text{if sick}}. \tag{4}$$

We do not model consumers' sensitivity to out-of-pocket prescription drug prices when choosing between pharmacies. Drug co-payments typically depend on the consumer's health insurance status or plan but are unlikely to vary substantially across pharmacies.[4]

Sick consumers demand both retail and medication products and solve

$$\max\left\{ \max_{m,j} \left\{ \underbrace{\left(\delta^M_{imt} + \delta^R_{ijt}\right)}_{\text{joint utility}} + \underbrace{\left(\overbrace{\gamma^M_i d_{im} + \gamma^R d_{ij}}^{\text{distaste for distance}} \times \overbrace{\mathbb{1}\{j \neq m\}}^{\text{foot traffic spillover}} \right)}_{\text{distaste from travel}} + \underbrace{\epsilon_{imjt}}_{\text{taste shifter}} \right\}, 0 \right\} \tag{5}$$

where $d_{ij} := ||\ell(i) - \ell(j)||$ is the distance from consumer i to store j. We introduce random coefficients γ^M_i on distance to capture heterogeneity in urgency that might arise from varying medical conditions. The spillover from pharmacy business to retail sector is captured by the indicator function $\mathbb{1}\{j \neq m\}$, which allows consumers to forgo the travel cost when purchasing retail products and filling prescriptions at the same store.

Healthy consumers do not have demand for prescription drugs and simply choose retailer j by solving

$$\max\left\{ \max_{j} \left\{ \delta^R_{ijt} + \gamma^R d_{ij} + \epsilon_{ijt} \right\}, 0 \right\}. \tag{6}$$

Supply. Pharmacy m's profit, given its location $\ell(m)$, is given by

$$\pi_{mt}\left(p^R_{mt}; \ell(m)\right) = \underbrace{\overbrace{\left(p^M_{mt} - c^M_{mt}\right)}^{\text{exogenous}} q^M_{mt}}_{\text{medication profit}} + \underbrace{\left(p^R_{mt} - c^R_{mt}\right) q^R_{mt}}_{\text{retail profit}} - \underbrace{\phi_{\ell(m)t}}_{\text{fixed cost}} + \underbrace{\varepsilon_{mt}}_{\text{idio. shock}}, \tag{7}$$

[4] We can in principle include the out-of-pocket payment in the utility equation because we are able to observe this payment in the claims data. This will allow us to convert medication-sector utility into dollar equivalent units, but this assumption will not influence our demand estimation as long as the out-of-pocket payment for a given consumer does not vary by pharmacy.

Table 2: Identification

Parameters	Source of identification	Data source
A. Common preferences		
price sensitivity α_i	Variation in prices	Nielsen scanner + Safegraph spend
B. Retail sector		
distaste for distance γ^R	Variation in retail distance	Nielsen scanner + Safegraph visitor location
unobservable quality ξ^R_{jt}	Market shares	Safegraph visitor location
C. Medication sector		
distaste for distance γ^M_i	Variation in pharmacy distance	APAC claims + Safegraph visitor location
unobservable quality ξ^M_{mt}	Market shares	APAC claims
D. Supply side		
cost of retail c^R_{mt}	Pharmacy first order conditions	
fixed cost $\phi_{\ell(m)t}$	Observed pharmacy locations	APAC claims + NPI data

where c^M_{mt} and c^R_{mt} are the marginal costs of medication and retail products, $\phi_{\ell(m)t}$ is the fixed cost, and ε_{mt} is an idiosyncratic shock. We will treat prescription drug margins $p^M_{mt} - c^M_{mt}$ in the medication sector as exogenous, taking drug reimbursement rates directly from the data, to avoid modeling the bargaining process with insurers. As a first pass, we will also abstract away from dynamics and avoid endogenizing pharmacy entry and location choices, instead focusing on the profitability of existing pharmacies under various counterfactuals.[5]

4.2. **Identification**

We informally summarize the high-level intuition behind our identification strategy in Table 2, where we list the parameters to estimate and potential sources of variation. The fixed costs $\phi_{\ell(m)t}$ will rationalize the observed spatial distribution of pharmacies. One concern is that we do not observe the total front-end retail market size, which we may need to calibrate using aggregate moments such as the revenue share of front-end stores.[6]

5. Counterfactuals and next steps

Our immediate next steps are to finish the application process for the Oregon APAC data, build upon our existing descriptive evidence, and obtain preliminary demand estimates. We also outline some ideas for potential counterfactuals to study.

[5] Our goal for now is to keep the framework simple by avoiding a high-dimensional, combinatorial entry problem. In counterfactual analyses, we can still analyze pharmacy closure by taking the existing pharmacies from the data.

[6] For example, we know that the front-end retail sector accounts for 25% of the total revenue at CVS and Walgreens nationwide from their annual reports.

Quantifying the effects of cross-subsidization. Our first counterfactual exercise analyzes the impacts of cross-subsidization and cross-market spillovers on inequality and efficiency. We consider two layers of counterfactuals: (i) shutting down the foot-traffic spillover term in consumer problem (5) and (ii) banning pharmacies from operating retail sectors. Under these counterfactuals, we are interested in analyzing profits for existing pharmacies—in particular, how the change in profits may be correlated with income across space—which could provide a rough approximation to potential exit when the cross-subsidization channel is limited. We view these counterfactuals as a quantification exercise for our cross-subsidization mechanism and a demonstration of how this mechanism may interact with spatial inequality as opposed to a concrete, policy-relevant proposal.

Changes to prescription drug reimbursement rates. According to a National Community Pharmacists Association (NCPA) survey, one of the main contributing factors toward pharmacy deserts are low prescription drug reimbursement rates.[7] It is also commonly argued that Medicaid reimbursement rates are especially low, disincentivizing pharmacies from operating in low-income communities. We can simulate the effects of adjusting Medicaid reimbursement rates on pharmacy profits, taking into account pharmacies' joint profits from both retail and medication sectors, using our estimated model.

Pharmacy-retailer mergers. Our theory of cross-subsidization and cross-market spillovers rationalizes mergers and acquisitions between pharmacies and retailers. A recent example is CVS Health's acquisition of Target's pharmacy business. We can evaluate the effects of these mergers and acquisitions by simulating counterfactuals where CVS and Target, or other joint pharmacy and retail entities, are forced to be separate.

Oregon's Corporate Activity Tax. We focus on the state of Oregon, where we have comprehensive pharmacy claims data, to estimate our model. Beginning in tax year 2020, Oregon implemented a Corporate Activity Tax that taxes companies on "corporate acitivity" above $1 million. This tax, which effectively applies to revenues rather than profits, is cited as a major reason for the closure of nearly 60 Bi-Mart pharmacies in 2021.[8] With our estimated supply-side model, we can study the effects of the Corporate Activity Tax on pharmacy access.

[7] See https://www.ncpa.co/pdf/survey-health-cp.pdf.

[8] According to Bi-Mart spokesman Don Leber, pharmacies have high revenues, but when they subtract the high costs of drugs and overhead, they end up with low profits.

References

Allcott, H., R. Diamond, J.-P. Dubé, J. Handbury, I. Rahkovsky, and M. Schnell (2019): "Food deserts and the causes of nutritional inequality," *The Quarterly Journal of Economics*, 134, 1793–1844.

Faulhaber, G. R. (1975): "Cross-subsidization: pricing in public enterprises," *The American Economic Review*, 65, 966–977.

Gannaway, G. (2016): "Pharmacy Deserts and Medication Adherence," *Working Paper*.

Kim, H.-J. (2023): "Rural Pharmacy Access and Competition: Static Games with Machine Learning," *Working Paper*.

Lindrooth, R., M. Perraillon, R. Hardy, and G. Tung (2018): "Understanding The Relationship Between Medicaid Expansions And Hospital Closures," *Health Affairs*, 37.

Moghtaderi, A., J. Pines, M. Zocchi, and B. Black (2020): "The effect of Affordable Care Act Medicaid expansion on hospital revenue," *Health Economics*, 29, 1682–1704.

Qato, D. M., M. L. Daviglus, J. Wilder, T. Lee, D. Qato, and B. Lambert (2014): "'Pharmacy deserts' are prevalent in Chicago's predominantly minority communities, raising medication access concerns," *Health Affairs*, 33, 1958–1965.

Shoag, D. and S. Veuger (2018): "Shops and the city: Evidence on local externalities and local government policy from big-box bankruptcies," *Review of Economics and Statistics*, 100, 440–453.

Thomassen, Ø., H. Smith, S. Seiler, and P. Schiraldi (2017): "Multi-category competition and market power: a model of supermarket pricing," *American Economic Review*, 107, 2308–2351.

Wu, R. (2024): "Pharmacy Switching Cost and Retail Traffic Growth," *Working Paper*.

From: **Institutional Review Board for Human Subjects** irb@Princeton.EDU
Subject: RE: IRB waiver for pharmacy desert study
Date: April 16, 2024 at 8:36 AM
To: Renjie Bao renjie.bao@princeton.edu
Cc: Ranie Lin ranielin@princeton.edu

IS

On April 16, 2024, the IRB determined that the proposed activity is not human subjects research as defined by DHHS and FDA regulations. Consequently, Princeton IRB approval is not applicable. You are welcome to pursue the activity, obtaining any applicable administrative or departmental (non-IRB) approvals.

From: Renjie Bao <renjie.bao@princeton.edu>
Sent: Tuesday, April 16, 2024 8:00 AM
To: Institutional Review Board for Human Subjects <irb@Princeton.EDU>
Cc: Ranie Lin <ranielin@princeton.edu>
Subject: Re: IRB waiver for pharmacy desert study

Thanks for providing these information!! After consulting with the data provider, we understand that it is infeasible for us to get the individual identifier. We hereby confirm that our research is limited to data analysis, we are not involved in the collection, and the data is not identifiable. Could you help us generate an IRB waiver?

Sorry for any inconvenience and thank you for the assistance!!

Best,
Renjie

On Apr 15, 2024, at 2:11 PM, Institutional Review Board for Human Subjects <irb@Princeton.EDU> wrote:

Thanks for the clarification. Since the data is identifiable, the activity is human subjects research. Please submit an IRB application. Below and attached are resources.

https://ria.princeton.edu/human-research-protection/resources-and-quick-links/ohrp-frequently-asked-que-1

We review minimal risk studies within five business days of submission. After the study is approved, you can implement it.

If you have IRB questions, let me know. The below link answers most eRIA/tech questions.

https://ria.princeton.edu/eRIA/eRIA-Help-Training

If you experience tech issues, the tech team can help: eria-irb@princeton.edu. Thanks.

From: Renjie Bao <renjie.bao@princeton.edu>
Sent: Monday, April 15, 2024 1:34 PM
To: Institutional Review Board for Human Subjects <irb@Princeton.EDU>
Cc: Ranie Lin <ranielin@princeton.edu>
Subject: Re: IRB waiver for pharmacy desert study

Hello,

I hereby confirm that our research is limited to data analysis and we are not involved in the collection. However, the data we are looking for contains identifiable information up to street address of each individual. Could you please advise whether we are qualified for IRB waiver?

Thanks,
Renjie

> On Apr 15, 2024, at 9:01 AM, Institutional Review Board for Human Subjects <irb@Princeton.EDU> wrote:
>
> Hello,
>
> Please confirm the following information:
>
> - The activity is limited to data analysis.

- You were not involved in the collection.
- The data is not identifiable. The regulations define “identifiable” as *information for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information.* Note that “identifiable” means identifiable to the investigator, not identifiable in recordings or publications.

From: Renjie Bao <renjie.bao@princeton.edu>
Sent: Monday, April 15, 2024 8:19 AM
To: Institutional Review Board for Human Subjects <irb@Princeton.EDU>
Cc: Ranie Lin <ranielin@princeton.edu>
Subject: IRB waiver for pharmacy desert study

Dear IRB,

We are Ranie and Renjie, PhD students from Economics department. We are working on a project studying pharmacy deserts and are preparing to apply for access to the Oregon All-Payer All-Claims (APAC) database.

Our research will use purely secondary data but we plan to apply for APAC data elements including member zip code or member street address. Could you please advise us on our eligibility for an IRB waiver?

Thank you for your guidance and support!

Best,
Renjie

<NewStudy.pdf>

From: Ranie Lin
To: OLIVER James
Cc: OHPR - APAC Admin; Piper Block (she/her); renjie.bao@princeton.edu
Subject: Re: APAC data request 6304
Date: Monday, October 14, 2024 2:56:40 PM

Hi James,

Our IRB granted us a waiver after we contacted them and were asked to confirm the following details:
- The activity is limited to data analysis
- We were not involved in data collection
- The data is not identifiable, defined as "information for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information."

It was our understanding that we would not be able to recover the identities of specific subjects from the data that we are requesting.

Ranie

On Mon, Oct 14, 2024 at 5:48 PM OLIVER James <JAMES.OLIVER@oha.oregon.gov> wrote:

> Did you submit an application to your IRB? If so, please send a copy. If not, please explain how the waiver was granted.
>
> Thank you.
>
> **James Oliver**
>
> Office of Health Analytics
>
> Oregon Health Authority
>
> james.oliver@oha.oregon.gov
>
> 
>

From: Ranie Lin <ranie.lin@gmail.com>
Sent: Monday, October 14, 2024 2:43 PM
To: OLIVER James <JAMES.OLIVER@oha.oregon.gov>
Cc: OHPR - APAC Admin <APAC.Admin@odhsoha.oregon.gov>; Piper Block (she/her) <Piper.Block@oha.oregon.gov>; renjie.bao@princeton.edu
Subject: Re: APAC data request 6304

Hi James,

We did not submit any additional materials to our IRB -- they granted us a waiver because we will not have direct identifiers in our data. Could you let us know if there's something else we need to submit?

Thank you,

Ranie

On Mon, Oct 14, 2024 at 4:19 PM OLIVER James <JAMES.OLIVER@oha.oregon.gov> wrote:

> Mr. Lin:
>
> Please forward a copy of the research protocol that was submitted to your IRB. This will allow us to proceed with your application.
>
> Thank you.
>
> **James Oliver**
>
> Office of Health Analytics

Oregon Health Authority

james.oliver@oha.oregon.gov

Security Risk Management Plan

Custodian: Ranie Lin (PI)

We will store and analyze the APAC data on the Citadel system, which is a secure and compliant computing environment sponsored by Princeton University. The Citadel system, located at Princeton University's High-Performance Computing Research Center, is designed to meet stringent security requirements, making it an ideal platform for handling APAC data. This document outlines how Citadel addresses secure storage, access restrictions, user account controls, personnel training, data derivative protections, direct identifier handling, breach management, compliance with HIPAA and HITECH, and electronic device protections to ensure the utmost security of APAC data.

Secure Storage

Citadel ensures the secure storage of APAC data through advanced encryption techniques. Electronic data is encrypted both at rest and in transit using FIPS 140-2 compliant standards. Hard copies, if needed, will be stored in physically secured locations with controlled access. Citadel's tiCrypt client provides remote desktop access to Secure Virtual Machines (SVMs), ensuring that data is processed within an isolated and encrypted environment.

Access Restrictions

Access to APAC data within Citadel is strictly controlled and limited to individuals listed on the data use agreement. Role-based access controls (RBAC) are enforced, and data export is restricted to authorized individuals only. The tiCrypt client ensures that only authenticated users can interact with the SVMs, and all access attempts are logged for audit purposes.

User Account Controls

Citadel implements robust user account controls, including:

1. Password Protections: Strong password policies are enforced using the zxcvbn library, requiring high complexity and regular updates.
2. Maximum Failed Login Attempts: Accounts are locked after five failed login attempts to prevent unauthorized access.

3. Idle Lockout Periods: Sessions lock after five minutes of inactivity, and users must re-authenticate to continue.
4. User Audit Logs: Detailed logs of user activities are maintained and regularly reviewed to detect and respond to unauthorized access attempts.

Personnel Training

All personnel accessing APAC data through Citadel are required to undergo training on managing protected health information (PHI) and other sensitive data. This training covers data handling procedures, security protocols, and compliance requirements, with annual refresher courses to ensure ongoing awareness and adherence to best practices.

Protection of Data Derivatives

Derivatives of APAC data that are re-identifiable are treated with the same level of security as the original data. Citadel ensures that any analysis or reports generated are encrypted and access-restricted. The tiCrypt system manages metadata and auxiliary data securely, preventing re-identification.

Handling Direct Identifiers

We do not inquire direct identifiers from the APAC.

Breach Management

In the event of a data breach, Citadel follows comprehensive procedures:

1. Identification: Immediate identification and assessment of the breach.
2. Reporting: Timely notification of relevant stakeholders, including regulatory bodies.
3. Remediation: Rapid containment and mitigation measures, including isolating affected systems and conducting forensic analysis.
4. Review and Update: Post-incident reviews to update security protocols and prevent future breaches.

Compliance with HIPAA and HITECH

Citadel's security measures comply with the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) Act. Citadel ensures the confidentiality, integrity, and availability of protected health information (PHI) through rigorous controls and protocols that meet or exceed regulatory requirements.

Electronic Device Protections

Citadel provides extensive protections for electronic devices accessing APAC data:

1. Anti-Virus/Anti-Malware Software: Devices are equipped with up-to-date anti-virus and anti-malware software, and VM images are scanned regularly.
2. Firewalls: Network firewalls prevent unauthorized access and monitor traffic.
3. Network Encryption: Data transmitted over networks is encrypted using SSL/TLS protocols, ensuring secure communication channels.

Conclusion

Citadel's comprehensive security framework ensures that APAC data is protected at every stage of storage, access, processing, and transmission. By implementing stringent access controls, user account protections, personnel training, data derivative safeguards, direct identifier handling, breach management, compliance with **HIPAA** and **HITECH**, and robust electronic device protections, Citadel provides a secure and compliant environment for handling sensitive APAC data.

Please answer each of the following questions:

Please indicate the year(s) of data requested	2019 - 2022
Do you want people who are not Oregon residents and their claims included? People with Medicaid coverage or Medicare coverage reported by CMS are Oregon residents regardless of address.	Yes x
Do you want people with pharmacy coverage, but no medical coverage or claims included?	Yes x
Do you want people with dental coverage, but no medical coverage or claims included?	No x
Do you want orphan claims included? (claims, but no eligibility or coverage reported)	No x
Do you want denied claims included? (No reason is provided for denied medical or pharmacy claims. Claims can be denied then paid)	Yes x
Do you want pharmacy claims for people with pharmacy coverage, but no medical coverage or claims included?	Yes x
Do you want dental claims for people with dental coverage, but no medical coverage or claims included?	No x

What payer types do you want?	Commercial, Medicaid, Medicare Adv.
	x

One payer reported the claim status for all of their claims as fee-for-service for some years when most claims were encounter or managed care claims. Do you want the claim status changed to managed care?	Change to encounter
	x

What medical claim types do you want?	Other
	x

Do you want to limit medical claims data to selected diagnoses, procedure or other codes?	No
	x

Do you want substance use disorder claims (SUD)? SUD claims were not available for request prior to APAC release 14. SUD requests require detailed information about purpose, hypotheses and analyses, information about data access, security, data destruction and data linking to any other source and detailed justification for requested data elements. Date use and release of information are restricted. Requires additional Data Use Agreement	No
	x

Do you want Coordination of Benefit (COB) medical claims?	No
	x

Do you want pharmacy claims?	Yes
	x

Do you want pharmacy claims for people with pharmacy coverage, but no medical coverage or claims included?	Yes
	x

Do you want Coordination of Benefit (COB) pharmacy claims?	Yes, when both the primary and secondary payer report
	X

Do you want dental claims?	No
	x

Do you want dental claims for people with dental coverage, but no medical coverage or claims included?	No
	x

Do you want monthly eligibility data (insured/covered by month, by payer, by plan)?	Yes
	x

Do you want claims and eligibility data for selected age groups only?	All ages
	x

Do you want to limit claims and eligibility data by sex/gender?	Include all
	x

Are you requesting identifiable data?	Zip code	County
	x	x

Do you want provider data?	Yes
	x

Do you want APAC data linked to Oregon Center for Health Statitistics (CHS) Death Certificate data and/or Birth Certificate data? Please include a list of the birth and or death data variables that you plan to request from birth and/or death certificate data. You will need approval from both CHS and APAC. Submit request to APAC first. After APAC approval submit request to CHS and provide APAC approval notice. https://www.oregon.gov/oha/PH/BIRTHDEATHCERTIFICATES/VITALSTATISTICS/Pages/Data-Use-Requests.aspx	No
	x

Is your requested APAC data going to be linked by the APAC Team or data requester to any other data source?	No
	x

Field Requested	Data Element	Security Level	Description	Justification
The data elements highlighted in blue are provided in every data request	uid	De-Identified	A unique identifier that links to the row as submitted in the PC Intake File Layout. Used for linking tables/views	
	release_id	De-Identified	A value associated with the data release	
	dw_claim_id	De-Identified	A unique medical claim identifier	
	pc032_prescription_fill_dt	De-Identified	Prescription fill date	
	dw_member_id	De-Identified	A payer & plan specific unique identifier for a person. A person can have multiple member IDs for a single payer because they can have multiple plans. DW_member_IDs are not unique identifiers for a person across payers and years	
	uniquepersonID	De-Identified	A unique identifier for a person across payers and time	
	dw_person_id	De-Identified	Vendor identifier for a person across payers and time-2 million people assigned more than one identifier	
	pc025_claim_status_cd	De-Identified	Claim status. P (Paid), D (Denied), C - (MCO/CCO encounter) E (other)	
	pc003_insurance_product_type_cd	De-Identified	A code that indicates an insurance coverage type	
	orphan_fl	De-Identified	Identifies orphan claim with no corresponding eligibility for the date of service. 1 (Yes), 0 (No)	

	member_state	De-Identified	People with Medicaid coverage and people with Medicare coverage reported by the Centers for Medicare & Medicaid Services are Oregon residents regardless of reported address	
	Suppressed_FI	De-Identified	1 (denied claim line), 0 (other than denied)	
	RemovedReversal_FI	De-Identified	1 (claims not included before release 13 because the charge, paid amount, and allowed amounts are zero or zero when summed across claim lines and after the removal of denied claim lines, 0 (otherwise)	
x	pc025_claim_status_cd	De-Identified	Claim status. P - Paid,C - CCO encounter, E - other	Basic information on claim status is needed for understanding which
x	COBDup	De-Identified	Links claims based on uniquepersonID, date, pc_026_drug_cd, charged amount, and provider and identifies an event that could be either COB claim or duplicate paid claim	Identifying COB or duplicate paid claims is needed to ensure that we can correctly identify unique payment events without double counting.

x	pc001_payer_type	De-Identified	Payer reported payer type codes:(C) Carrier, (D) Medicaid, (G) Other government agency, (P) Pharmacy benefits manager, (T) Third-party administrator, (U) Unlicensed entity	The type of payer (e.g., carrier, Medicaid, etc…) is needed for us to create a broad classification of claims and assess how different types of insurance coverage affect access to pharmacy services across different regions and socio-economic groups.
x	Claim_LOB	De-Identified	Payer line of business: 1 (Medicare), 2 (Medicaid), 3 (commercial, 0 (no line of business reported)	Similar to the payer type, the payer line of business (e.g., Medicare, Medicaid, etc…) is needed for us to create a broad classification of claims and assess how different types of insurance coverage affect access to pharmacy services across different regions and socio-economic groups.
x	self_insured_fl	De-Identified	Self Insured flag	A flag indicating whether the claim is from a self-insured plan will allow us to understand the role of employer-provided healthcare in pharmacy access and utilization patterns.

x	dw_pharmacy_id	De-Identified	A unique identifier associated with a unique pharmacy across plans, payers and years	Identifier of the pharmacy is needed to link with other pharmacy details that we are requesting.
x	dw_prescribing_provider_id	De-Identified	A unique identifier associated with a unique prescribing provider across plans, payers and years	Identifier of the provider is needed to link with provider composite data.
x	pc021_pharmacy_npi	De-Identified	Pharmacy's National Provider Identifier (NPI)	The National Provider Identifier for pharmacies enables precise tracking of pharmacy service availability and is crucial for geospatial analysis of pharmacy deserts.
x	pc021a_pharmacy_alt_id	De-Identified	Pharmacy's alternate identifier as assigned by the payer	An alternative identifier for pharmacies, if available, will be used for data validation or cross-referencing.
x	pc020_pharmacy_name	De-Identified	Name of pharmacy	Important pharmacy information to build geographical measure of pharmacy deserts.
x	pc022_pharmacy_city	De-Identified	City of pharmacy	Important pharmacy information to build geographical measure of pharmacy deserts.
x	pc023_pharmacy_state	De-Identified	State of Pharmacy	Important pharmacy information to build geographical measure of pharmacy deserts.

x	pc024_pharmacy_zip	De-Identified	Zip Code of Pharmacy	Important pharmacy information to build geographical measure of pharmacy deserts.
x	pc048_prescribing_physician_npi	De-Identified	Identifier for the provider who prescribed the medication as assigned by the reporting entity	The physician NPI will allow us to identify distance from the prescribing provider to the pharmacy, which will be an alternative measure of pharmacy access.
x	pc026_drug_cd	De-Identified	National Drug Code (NDC)	This code is essential for identifying the specific drugs dispensed, which is crucial for analyzing medication adherence patterns across different drug types and the availability of essential medications in different regions.
x	pc033_dispensed_qty	De-Identified	Quantity dispensed	The quantity of drug dispensed will allow us to impute medication costs on the pharmacy side, which will be important for us to estimate the profitability on the pharmacy level. It will also help supplement our estimates of medication adherence.

x	pc028a_alt_refill_no	De-Identified	Alternate refill number	Needed for tracking refills accurately so that we can impute prescription adherence, which is an important health outcome related to pharmacy deserts.
x	pc034_days_supply_qty	De-Identified	Number of days that the drug will last if taken at the prescribed dose	Needed for tracking and imputing prescription adherence over time, which is an important health outcome related to pharmacy deserts.
x	pc030_dispense_as_written_cd	De-Identified	Dispense as written. Indicates if drug substitution authorized	This field could allow us to identify substitution from branded to generic drugs, providing greater accuracy into the exact drugs dispensed by pharmacies, which is needed for both medication adherence and pharmacy cost estimates.

x	pc028_calc_refill_no	De-Identified	Processor's count of times prescription refilled	The number of refills calculated by the processor is critical for assessing medication adherence and prescription refill behaviors, which are indicators of access to and continuity of care.
x	pc017_paid_dt	De-Identified	Prescription Payment date	This element helps construct timelines for access to medications and builds dynamic evidence for pharmacy deserts
x	pc036_paid_amt	De-Identified	Payment made by payer. Does not include expected copayment, coinsurance or deductible by the member 0 if amt=0, blank if missing	This set of elements about payments helps inform us about pharmacy revenues from dispensing drugs -- in particular, the amount received by the pharmacy from insurers -- which are essential for our estimation exercise on patient demand and pharmacy entry.

x	pc035_charge_amt	De-Identified	Payer reported charges or billed amount for the service 0 if amt=0, blank if missing	This set of elements about payments, in combination with the dispensing fee, helps inform us about total expected pharmacy revenues from dispensing drugs, which are essential for our estimation exercise on patient demand and pharmacy entry.
x	pc037_ingredient_cost_amt	De-Identified	Ingredient cost/list price 0 if amt=0, blank if missing	This set of elements about list prices helps to give us an approximation to the total costs for pharmacies to acquire drugs (which will be supplemented with other measures, e.g. NADAC from the CMS), which are essential for our estimation exercise on patient demand and pharmacy entry.

x	pc039_dispensing_fee_amt	De-Identified	Dispensing fee paid 0 if amt=0, blank if missing	This set of elements about dispensing fees, in combination with the charge amount, helps inform us about total expected pharmacy revenues from dispensing drugs, which are essential for our estimation exercise on patient demand and pharmacy entry.
x	pc040_copay_amt	De-Identified	Expected Co-payment by the member and $0 patientpaid	The copayment received by the pharmacy from members helps inform us about (i) pharmacy revenues from dispensing drugs and (ii) cost-sharing burdens that members face, which are crucial for our estimation exercise on patient demand and pharmacy entry.

x	pc041_coinsurance_amt	De-Identified	Expected Co-insurance by the member and $0 patientpaid	The co-insurance amount received by the pharmacy from members helps inform us about (i) pharmacy revenues from dispensing drugs and (ii) cost-sharing burdens that members face, which are crucial for our estimation exercise on patient demand and pharmacy entry.
x	pc042_deductible_amt	De-Identified	Expected Deductible by the member and $0 patientpaid	The deductible amount received by the pharmacy from members helps inform us about (i) pharmacy revenues from dispensing drugs and (ii) cost-sharing burdens that members face, which are crucial for our estimation exercise on patient demand and pharmacy entry.

x	patientpaid	De-Identified	Expected Patient paid amount. Amount patient paid when sum of copayment,coinsurance and deductible is less than the amount of pc043_patient_paid_amt reported	The total paid amount received by the pharmacy from patients helps inform us about (i) pharmacy revenues from dispensing drugs and (ii) cost-sharing burdens that members face, which are crucial for our estimation exercise on patient demand and pharmacy entry.

Field Requested	Data Element	Security Level	Description	Justification
The data elements highlighted in blue are provided in every data request	uid	De-Identified	A unique identifier that links to the row as submitted in the MM Intake File Layout.	
	release_id	De-Identified	A value associated with the data release	
	year_Eligibility	De-Identified	Year of eligibility	
	month_Eligibility	De-Identified	Month of eligibility	
	dw_member_id	De-Identified	A unique identifier associated with a single plan and payer and assigned to all eligibility and claims records associated with a given individual for that plan/payer. An individual can have multiple member ids for a payer because they can have multiple plans.	
	uniquepersonID	De-Identified	A unique identifier for a person across	
	dw_person_id	De-Identified	Vendor identifier for a person across payers and time-2 million people assigned	
	me003_insurance_product_type_cd	De-Identified	coverage type	
	me018_medical_coverage_flag	De-Identified	Medical Coverage Flag not required when	
	me019_prescription_drug_coverage_flag	De-Identified	Prescription Drug coverage flag	
	me207_dental_coverage_flag	De-Identified	month	
	member_state	De-Identified	People with Medicaid coverage and people with Medicare coverage reported by the Centers for Medicare & Medicaid	
x	PartD	De-Identified	Medicare pharmacy type: Medicare AdvantageRX or MedicareFFSRX (Medicare Fee-for-service)	An important control variable to identify differerent effects on pharmacy access from different types of

x	DualMedicareMedicaid	De-Identified	Dual or null when not dual	Individuals with both Medicare and Medicaid (dual-eligible) may have different access to pharmacy services than those with one form of coverage; understanding their distribution can illuminate disparities that
x	RXnomedicalMM	De-Identified	Pharmacy coverage and no medical coverage during same year, month: RXnoMedical or null	Patients with pharmay coverage but without medical coverage could be indicative of different health statuses or out-of-pocket medication expenses, a factor that could influence medication non-adherence and pharmacy access.
x	me201_medicare_coverage_flag	De-Identified	Type of Medicare coverage for Medicaid members only. A - Part A, B - Part B, AB - Parts A and B, C - Part C, D - Part D, CD - Part C and D, X - other, Z - none, not required when ME001=E	An important control variable to differentiate the effects on pharmacy access from different types of Medicare insurance.

x	me013_member_gender_cd	De-Identified	Member Gender:M (male), F (female), and U (unknown)	Gender data is important for analyzing differences in healthcare utilization, and it would be important for us to understand whether there are disparities across genders.
x	age	De-Identified	Member age in years calculated on the first day of the month	Age is a critical factor in healthcare needs and medication use; patterns in different age groups could signal areas at risk of becoming pharmacy deserts.
x	me203_metal_tier	De-Identified	Health benefit plan metal tier for qualified health plans (QHPs) and catastrophic plans as defined in the ACA:0 (Not a QHP or catastrophic plan), 1 (catastrophic), 2 (bronze), 3 (silver), 4 (gold), 5 (platinum)	Different tiers of insurance coverage can affect medication affordability and pharmacy accessibility, which could be an important indicator for inequality in pharmacy access.

x	me205_high_deductible_health_flag	De-Identified	High Deductible Health Plan Flag	High deductible plans can lead to higher out-of-pocket costs, potentially affecting pharmacy access, medication adherence, and the sustainability of pharmacies in certain areas.
x	me206_primary_insurance_ind	De-Identified	Flag indicates primary insurance	An indication of whether the plan is the patient's primary insurance is needed for us to correctly attribute observed outcomes to the correct plan type.
x	MCAID_cde_medicare_status	De-Identified	Medicare status reported for Medicaid recipients: MA (Part A only), MAB (Part A & B), MABD (Part A,B&D), MAD (Part A & D), MB (Part B only), MBD (Part B & D), MD (Part D only)	Understanding the Medicare status can highlight disparities in access to pharmacies among different population segments with different types of coverage.
x	MCAID_cde_enroll_recip_status	De-Identified	Medicaid enrollment status: managed care enrolled cap payment (1), managed care enrolled no cap payment (3), not managed care enrolled cap payment (5), fee for service (6) or null	Understanding the Medicaid enrollment status can highlight disparities in access to pharmacies among different population segments with different types of coverage.

x	rarestre	De-Identified	The rarest race-ethnicity identified for a person across payers and years (only one identified per person): (P) Native Hawaiian or Pacific Islander, (B) Black or African American, (I) American Indian or Alaskan Native, (A) Asian, (H) Hispanic or Latino, (W) White, (O) other and (noRE) no race-ethncity reported	We want to understand potential disparities in pharmacy access across different race-ethnicities, which is especially important for highlighting the equity concern on healthcare access.
x	re1_race_cd	De-Identified	All races reported by all payers for all years for a person: (P) Native Hawaiian or Pacific Islander, (B) Black or African American, (I) American Indian or Alaskan Native, (A) Asian, (W) White, (O) other, (U) unknown, (R) refused and null	We want to understand potential disparities in pharmacy access across different races, which is especially important for highlighting the equity concern on healthcare access.
x	re2_ethncity_cd	De-Identified	All ethnicities reported by all payers for all years for a person: (H) Hispanic), (O) Not Hispanic, (U) unknown, (R) refused and null	We want to understand potential disparities in pharmacy access across differente ethnicities, which is especially important for highlighting the equity concern on healthcare access.
Data elements that are frequently denied				

x	me017_member_zip	Limited	Zip code-from the date of eligibility	For our purpose, it is crucial to have a measure of how accessible pharmacies are to each individual patient, i.e., an approximate estimate of how far patients have to travel to get to certain pharmacies. The zip code of each member is therefore essential for us to build a baseline measure of pharmacy deserts. Our research question critically relies on this data element.

Field Requested	Data Element	Security Level	Description	Justification
Provided in every data request	release_id	De-Identified	A value associated with the data release	
x	national_provider_id	De-Identified	National Provider Identifier (NPI)	Pharmacy Claim data
x	Addr_Type	De-Identified	Address type of provider (B) Business, (L) Location, (S) Secondary Location, (I) Provider Index	Knowing the address of pharmacies is essential to build our measure of pharmacy access, i.e., how far patients must travel to get to the pharmacy
x	Addr_Street_1	De-Identified	Address of provider	Knowing the address of pharmacies is essential to build our measure of pharmacy access, i.e., how far patients must travel to get to the pharmacy
x	Addr_Street_2	De-Identified	Address 2 of provider	Knowing the address of pharmacies is essential to build our measure of pharmacy access, i.e., how far patients must travel to get to the pharmacy
x	Addr_City	De-Identified	City of Provider	Knowing the address of pharmacies is essential to build our measure of pharmacy access, i.e., how far patients must travel to get to the pharmacy

x	Addr_State	De-Identified	State of provider	Knowing the address of pharmacies is essential to build our measure of pharmacy access, i.e., how far patients must travel to get to the pharmacy
x	Addr_ZIP	De-Identified	ZIP Code of provider - may include non-US codes	Knowing the address of pharmacies is essential to build our measure of pharmacy access, i.e., how far patients must travel to get to the pharmacy

New or Amended APAC Data Request Review (custom or OHA Business Associate)

Staff Reviewer: Oliver

DRTS Number: 6304

Date review completed: 10/14/2024

	Yes	No	N/A	Need more information
Is this a new APAC request?	X			
New APAC Request (skip to next section if amendment request):				
1.1 Project staff contact information provided	X			
1.2 Project technical staff information provided				
2.1 Project summary provided with adequate detail to identify a specific unambiguous project	X			Analyze spatial inequality in access to pharmacies.
2.2 Research questions provided with adequate detail	X			The key research question: How does access to pharmacies differ across low and high income communities and rural versus urban areas?
2.3 Described planned products and reports derived from requested data	X			Academic paper
2.4 Project begin and end date provided	X			End 01/01/2027
2.5 Acknowledgement that APAC data cannot be reused beyond the DUA	X			
2.5 Acknowledgement that data cannot be shared beyond the DUA	X			
3.1ab Data request purpose box checked & description	X			Research
3.2 Checked box for level of data identifiers	X			Limited
3.3 IRB application, approval memo, end date	X			
4.1 Completed data elements workbook	X			
4.2 Adequately described how the data elements requested are the minimum necessary	X			Thorough justification provided for each data element. Requested pharmacy fields are reasonable given that pharmacy is the focus of the research. Only 14 enrollment fields requests.
5.1 Plan provided to prevent re-identification	X			
5.2ab Plan to link APAC data to other data source		X		
5.2c Requests OHA to link APAC to other data		X		
5.2d Detailed data linking plan provided			X	
5.3 Provided adequate description of data management, security and data destruction plan	X			
Passes Minimum Necessary Review				This request is for pharmacy, member months, and provider data for 2019-2022. Scope of request is very reasonable given the proposed research.

	Yes	No	N/A	Need more information
Recommend management approval	**X**			This is a legitimate research request.