Updates to Methodology for Estimating Expected Counts of Cases Diagnosed in Year 2020 in the Context of November 2022 Data Transmission to NCI SEER

Communication to SEER Research Registry Principal Investigators and Managers

September 27, 2022

Serban Negoita

Rocky Feuer

SEER Case Completeness

- **Case completeness – critical measure of data quality**
  - Case completeness standard stipulated in the registry contracts
  - Calculated for latest year of diagnosis included with each SEER submission
  - Registry completeness estimates are compared against program standards and results are included in the registry Data Quality Profile
  - Similar to previous years, SEER Completeness Standard for November submission is 98%
- **Completeness calculated as an Observed to Expected ratio**,
  - Reported as a percentage
  - Expected Count - denominator calculated annually by NCI and transmitted to registries before submission

Need To Adjust Expected Counts?

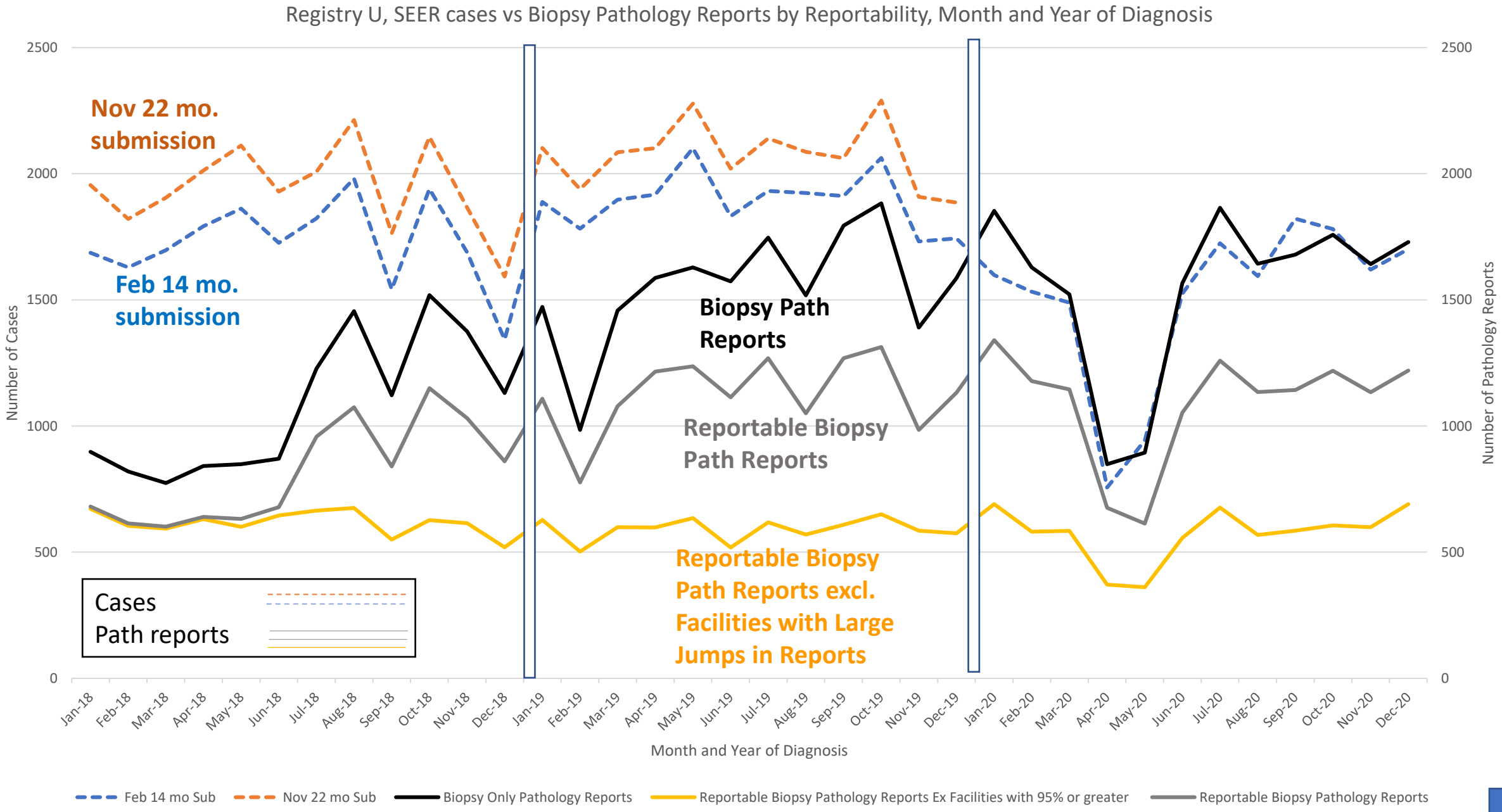
- **SEER uses joinpoint regression models to extrapolate expected count based on past submissions data**
  - Method referred to as the "internal method"
  - Likely to detect changes in registry operations, including operational delays
- **Projection assumes no sudden large changes in:**
  - the distribution of risk factors in underlying population (depending on the lag between changes in risk factors and the onset of cancer)
  - the patterns and methods of cancer early detection and diagnosis
- **Assumptions have been violated by COVID-19 pandemic**
  - changes in screening patterns
  - growth of telemedicine
  - population migration patterns

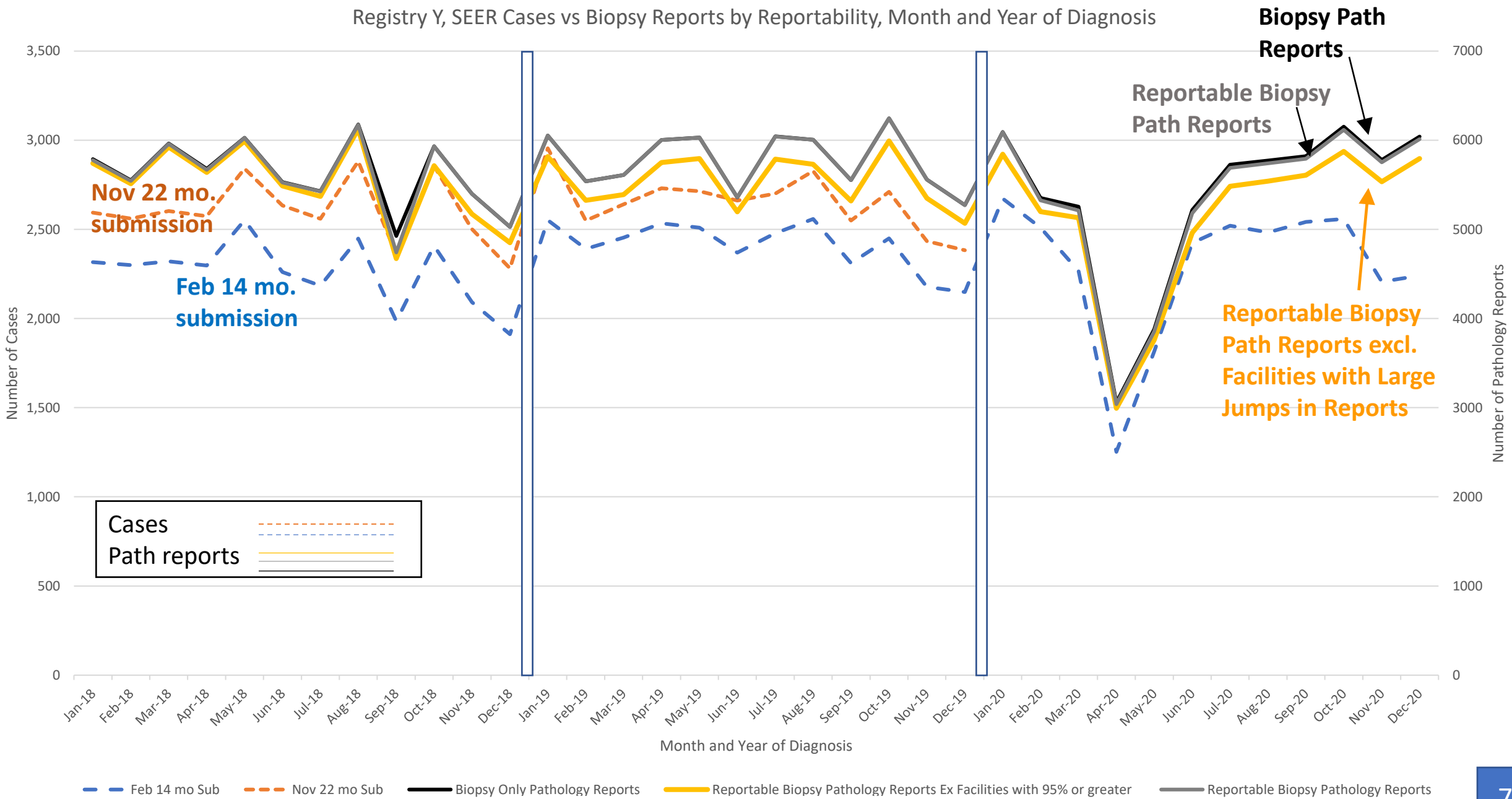
Data Sources Informing on True Changes in the Count of Incident Cases

- Lower number of pathology reports received by registries
  - Data available from Path Report Volume monitoring, including 12 SEER Registries
- Lower number of cases received with SEER February submission
- Lower number of cases received with NAACCR 12-mos. submission
- NAACCR and NPCR surveys
- Individual registry studies (e.g., NYSCR study using the hospital discharge data)
- Consultation with registry PIs and managers regarding operational delays vs. **true** decline in number of new cases

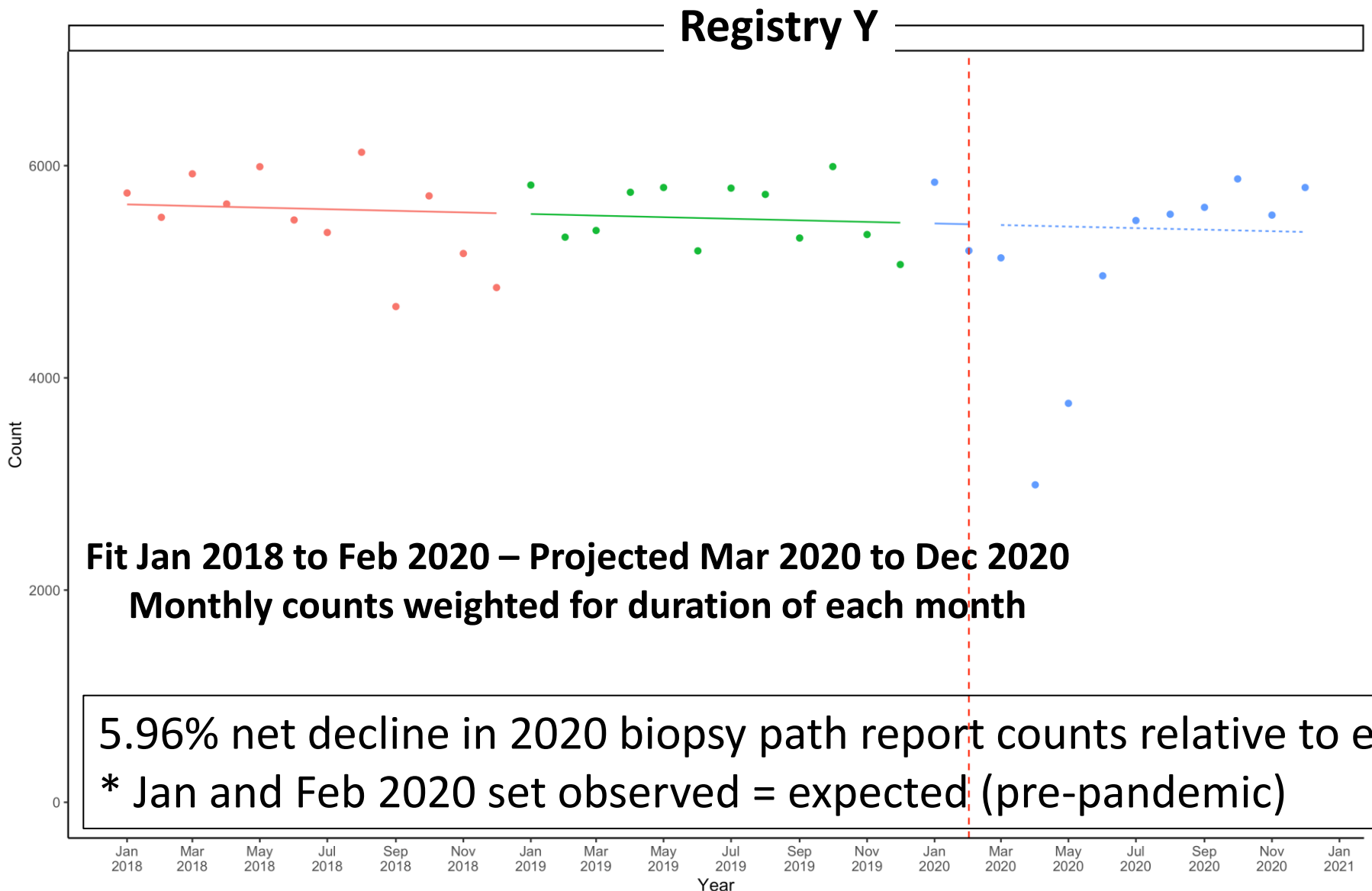
How to Quantify the True Decline in New Cases

- Find a data source:
  - Well correlated with new diagnoses
  - Not affected by operational delays
  - Consistently collected by most registries
  - Easily quantifiable
  - Counting would have minimal interference on registry operations
- Pathology reports satisfy most of the above criteria
  - Best option to adjust expected counts to reflect the true decline in new cases


Registry U, SEER cases vs Biopsy Pathology Reports by Reportability, Month and Year of Diagnosis
Nov 22 mo. submission
Feb 14 mo. submission
Biopsy Path Reports
Reportable Biopsy Path Reports
Reportable Biopsy Path Reports excl. Facilities with Large Jumps in Reports
Cases
Path reports
Number of Cases
Number of Pathology Reports
Month and Year of Diagnosis
Feb 14 mo Sub
Nov 22 mo Sub
Biopsy Only Pathology Reports
Reportable Biopsy Pathology Reports Ex Facilities with 95% or greater
Reportable Biopsy Pathology Reports


Registry Y, SEER Cases vs Biopsy Reports by Reportability, Month and Year of Diagnosis
Biopsy Path Reports
Reportable Biopsy Path Reports
Nov 22 mo. submission
Feb 14 mo. submission
Reportable Biopsy Path Reports excl. Facilities with Large Jumps in Reports
Cases
Path reports
Number of Cases
Number of Pathology Reports
Month and Year of Diagnosis
Feb 14 mo Sub
Nov 22 mo Sub
Biopsy Only Pathology Reports
Reportable Biopsy Pathology Reports Ex Facilities with 95% or greater
Reportable Biopsy Pathology Reports

Joinpoint Fit to Reportable Biopsy Path Reports excl. Facilities with >95% Jump



Fit Jan 2018 to Feb 2020 – Projected Mar 2020 to Dec 2020
Monthly counts weighted for duration of each month

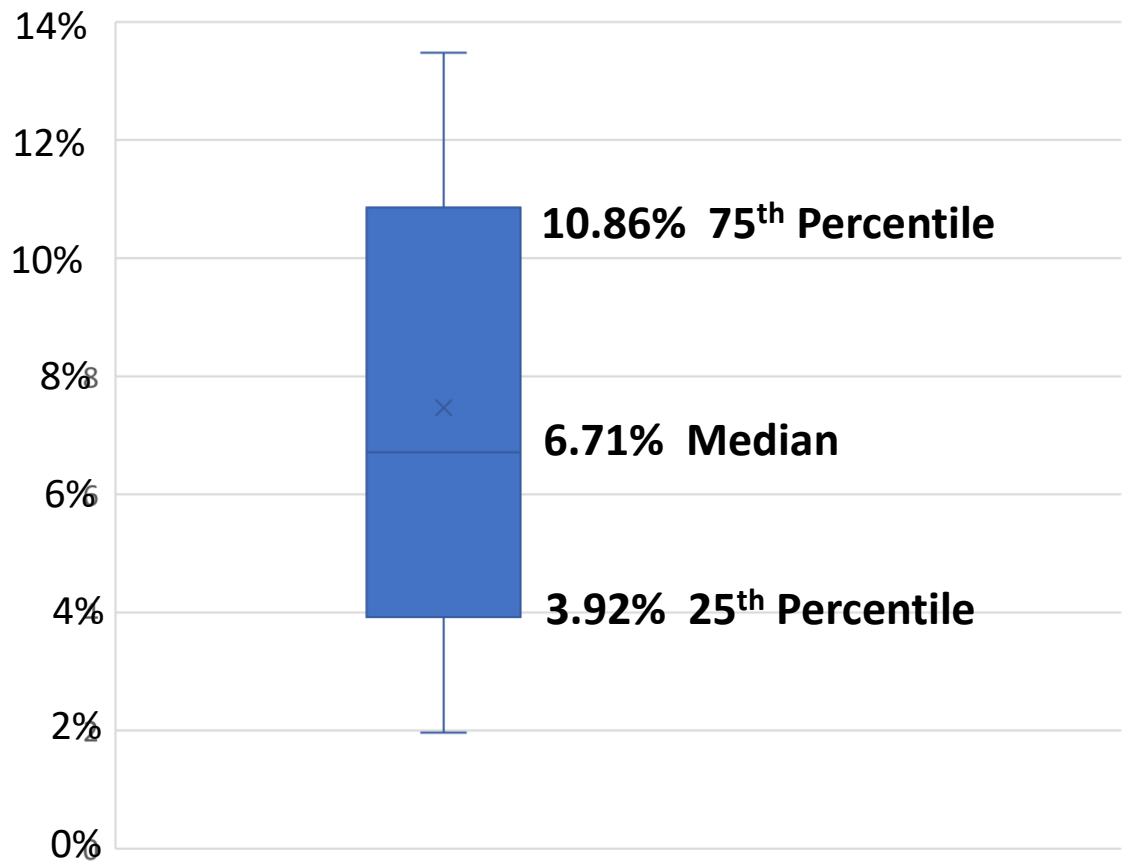
5.96% net decline in 2020 biopsy path report counts relative to expected*
* Jan and Feb 2020 set observed = expected (pre-pandemic)

Projected Completeness Estimation for November 2022 Submission Adjusted for Changes in E-Path Volume

	(Col F)	(Col G)		(Col M)	(Col N)		(Col S)	(Col T)	
	Feburary 2022 Completeness			Projected November Completeness Based on Delay Model			E-path Adjusted completeness, 11 registries		
Registry	Expected Count Nov 2022, SEER Internal Method	Observed Count, Feb 2022	Completeness 2022, SEER Internal Method (O/E)	Delay Factor (Ratio of Feb to November Delay Factor)	Estimated November 2022 Count (projected from Feb 2022 submission using delay) (col G) *(col M)	Projected Nov Completeness (col N)/ (col F)	Percent decline e-path report in 2020 due to pandemic	Adjusted Expected Count (Nov 2022) (1-col S)*(col F)	E-path-adjusted Estimated Nov 2022 Completeness (col N)/(col T)
Y	31,052	26,617	85.7%	1.092	29,061	93.6%	5.96%	29,201	99.5%

Eventual Cases = Feb Delay Factor * Feb Cases = Nov Delay Factor * Nov Cases
so
(Feb Delay Factor / Nov Delay Factor) * Feb Cases = Expected Nov Cases

Range of Percent Declines in 2020 e-Path Reports Among the 11 Registries that Participated



Limitations of Analysis

- In the process of honing the e-path data down to a reasonable analytic file we may have lost representativeness
- We only have e-path correction factors for 11 registries – we need factors for all registries
- In talking to registries recently, we found that most were already ahead of added cases between Feb and Nov as predicted by the delay model
  - This means that registries Nov counts will be larger than we computed based on the delay model, and thus their pandemic adjusted completeness should be better than we computed
- We can never know the counterfactual number of e-path reports that would have come in from March to Dec 2020 if the pandemic had never occurred
  - The projection is probably the most reasonable guess, but a 10-month projection is pretty long

Decision and Rationale

- We wanted to apply a single pandemic adjustment to all registries expected counts (in deference to the fact that we did not have adjustment factors for every registry)
- Given all of the vagaries of the analysis (as described on the last slide) we wanted to use a reasonably generous adjustment
- After extensive discussion among SRP leadership we decided on the 75th percentile of the distribution of computed % declines in e-path reports
  - This means that your expected counts for Nov will be reduced by 10.86%
  - The new expected numbers will be updated in SEER*DMS dashboard and in the calculation of SEER Edits by Oct. 1st at the latest.

Final Thoughts

- The nation is eagerly waiting for the release of 2020 cancer incidence rates.
- Researchers will carefully parse the drop in cases to help understand the impact of the pandemic on delays in the screening and diagnosis of cancer
  - They will implicitly assume that the decline is “real” rather than any unusual pandemic related issue with hospital or registry operations
- The pandemic adjusted expected cases may be overly generous to some registries, but we did not want the difficulty of computing the expected number in this very unusual year to be to the detriment to registries
- Please use the pandemic-adjusted expected count as your “base”, rather than a “goal” – we expect many registries to have completeness over 100%