Using health care data to study and improve colorectal cancer screening

3rd Seattle Symposium on Health Care Data Analytics

Jessica Chubak, PhD

October 23, 2018

Kaiser Permanente Washington Health Research Institute

KAISER PERMANENTE®

Disclosures and funding

- No Disclosures
- Funding (NIH):
  - R01CA121125 (SOS)
  - U24CA171524 (CRN)
  - U54CA163261, UM1CA222035 (PROSPR)

CRN CANCER RESEARCH NETWORK

PROSPR Population-based Research Optimizing Screening through Personalized Regimens

KAISER PERMANENTE®

Outline

- Part 1: CRC screening as a case study
- Part 2: Practical considerations
- Part 3: Summary and recommendations

Part 1

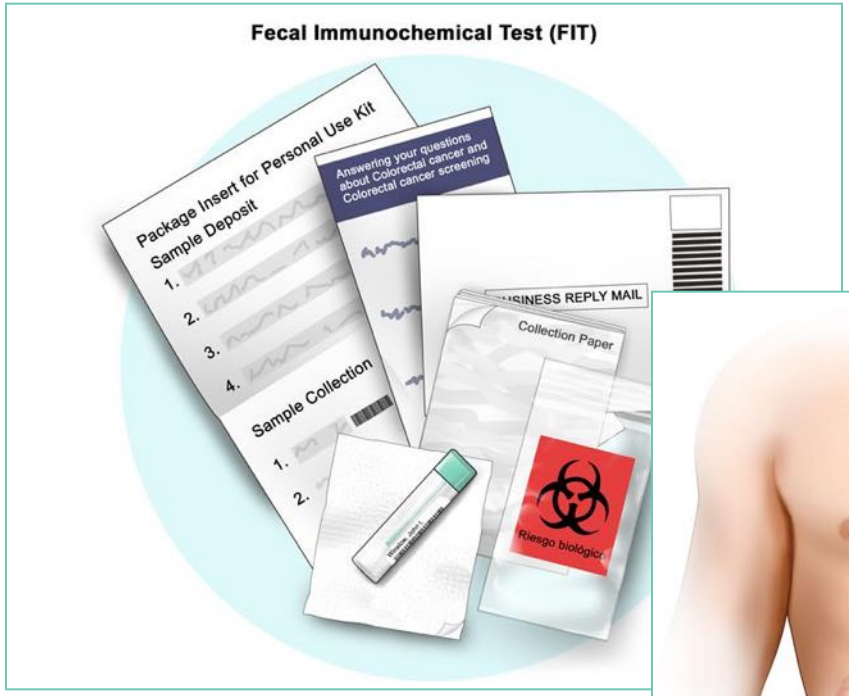
COLORECTAL CANCER SCREENING: A CASE STUDY

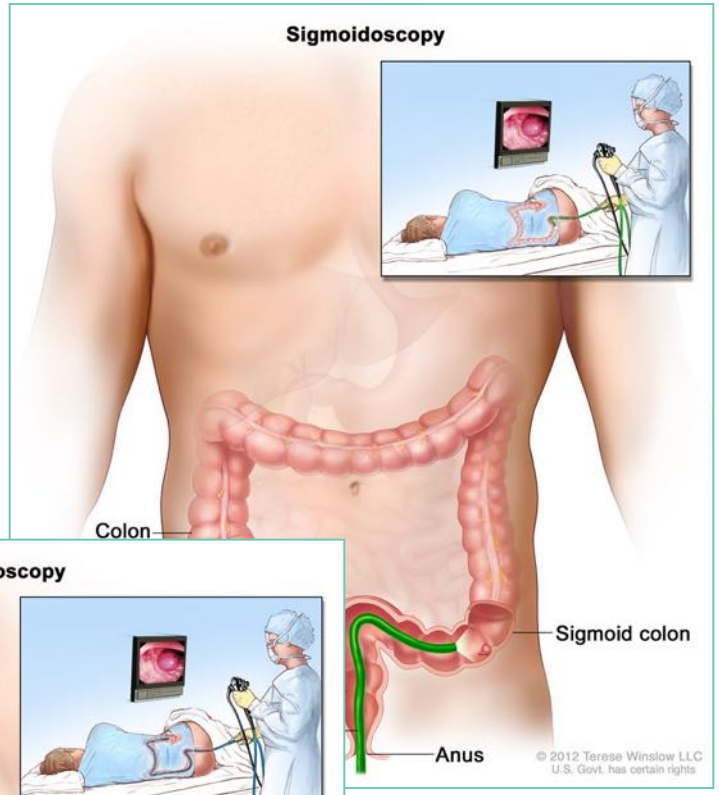
KAISER PERMANENTE®

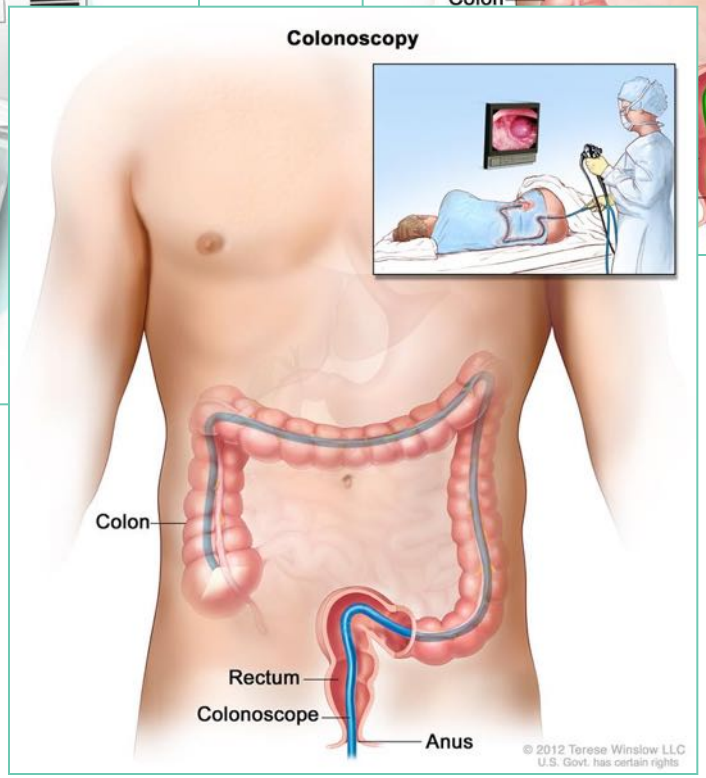
Why care about colorectal cancer screening?

- CRC is 3rd most common cancer in US
- Screening can prevent incidence and death
- Gaps in screening persist
- Comparative effectiveness of screening regimens is largely unknown

Types of CRC screening



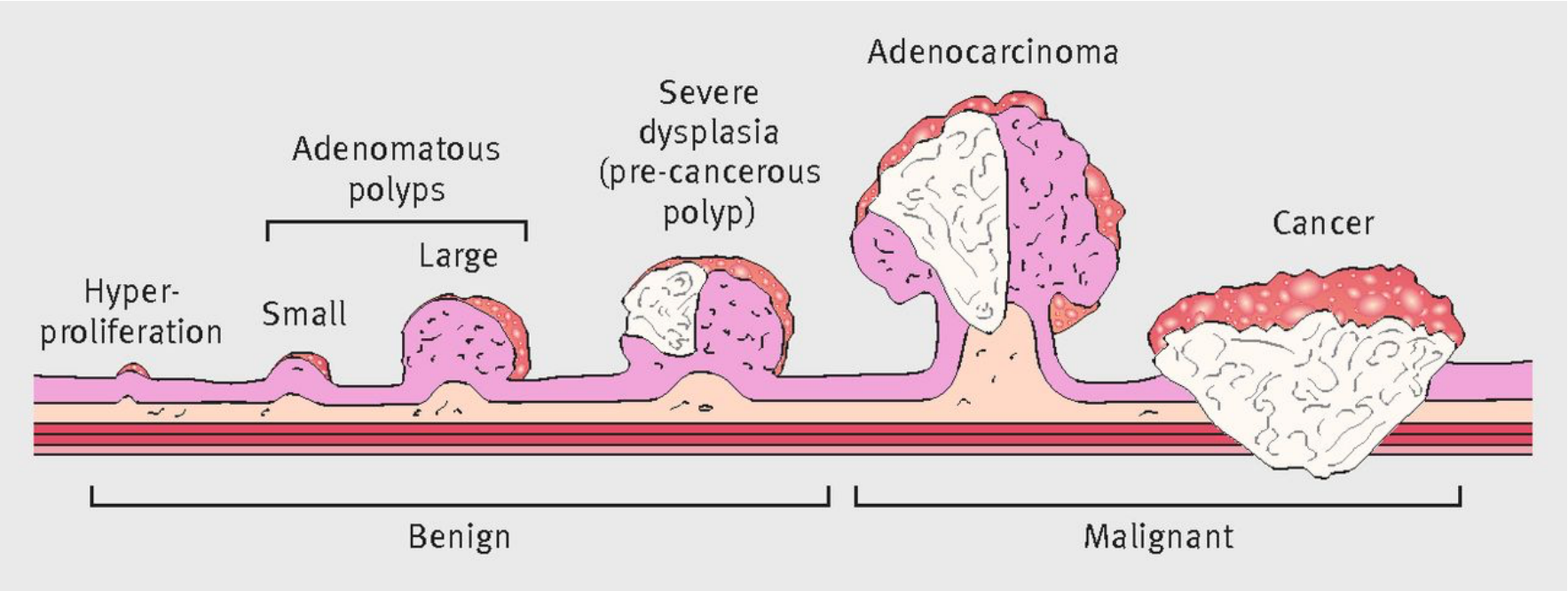




https://www.cancer.gov/types/colorectal/patient/colorectal-screening-pdq

KAISER PERMANENTE®

Adenoma – carcinoma sequence



Thrumurthy et al. BMJ 2016;354:i3590.

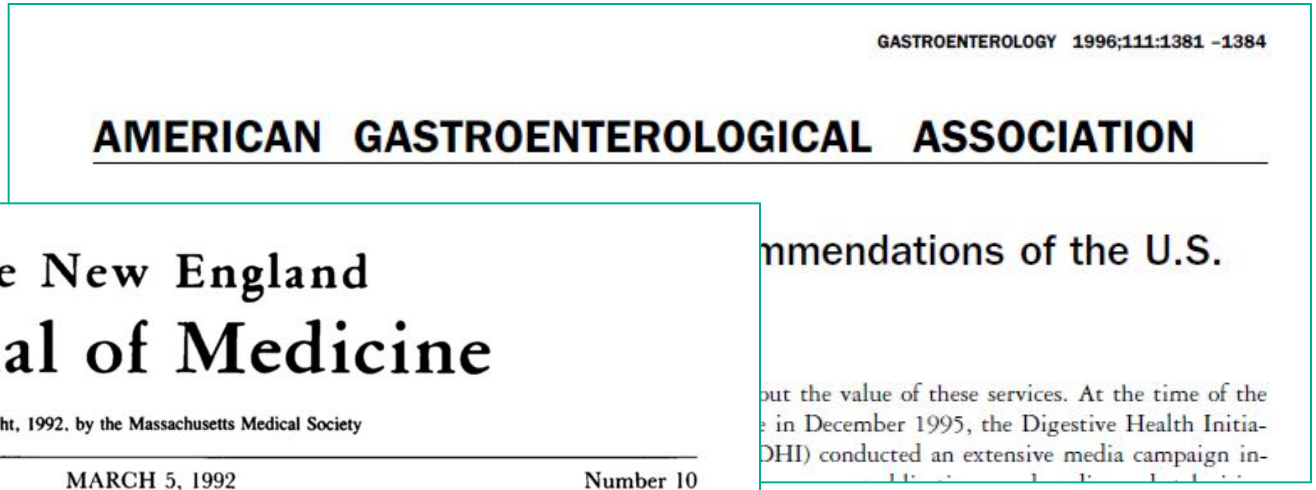
KAISER PERMANENTE®

Role of health care data in CRC screening

- What we **should** do
  - Evidence for guidelines
- What we **are** doing
  - Identification of care gaps
- What we **can** do
  - Foundation for interventions

Example 1: Evidence for guidelines
(What we *should* do)

Evidence → USPSTF Recommendations → Coverage

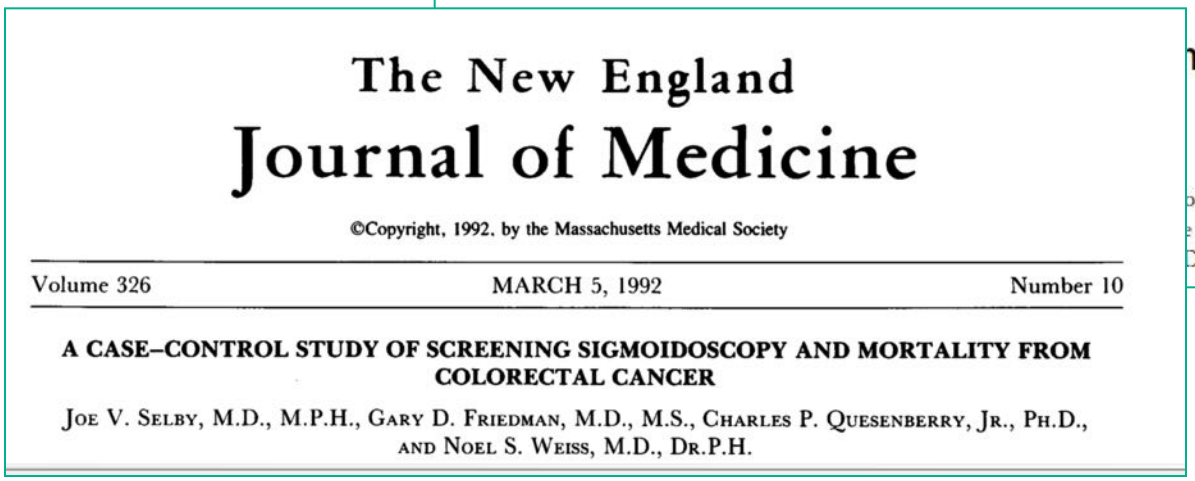


GASTROENTEROLOGY 1996;111:1381 –1384

AMERICAN GASTROENTEROLOGICAL ASSOCIATION

mmendations of the U.S.

ut the value of these services. At the time of the
in December 1995, the Digestive Health Initia-
DHI) conducted an extensive media campaign in-



The New England
Journal of Medicine

©Copyright, 1992, by the Massachusetts Medical Society

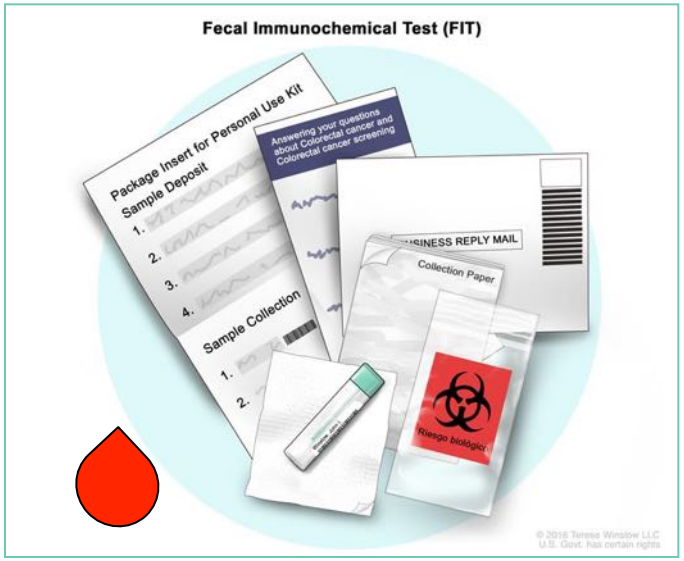
Volume 326 MARCH 5, 1992 Number 10

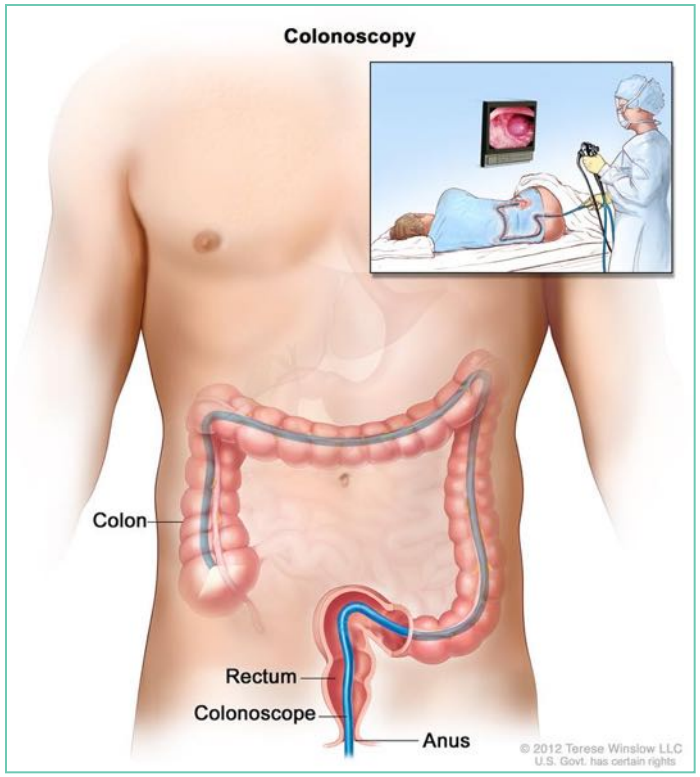
A CASE–CONTROL STUDY OF SCREENING SIGMOIDOSCOPY AND MORTALITY FROM COLORECTAL CANCER

JOE V. SELBY, M.D., M.P.H., GARY D. FRIEDMAN, M.D., M.S., CHARLES P. QUESENBERRY, JR., PH.D., AND NOEL S. WEISS, M.D., DR.P.H.

KAISER PERMANENTE®

Example 2: Identification of care gaps (What we *are* doing)



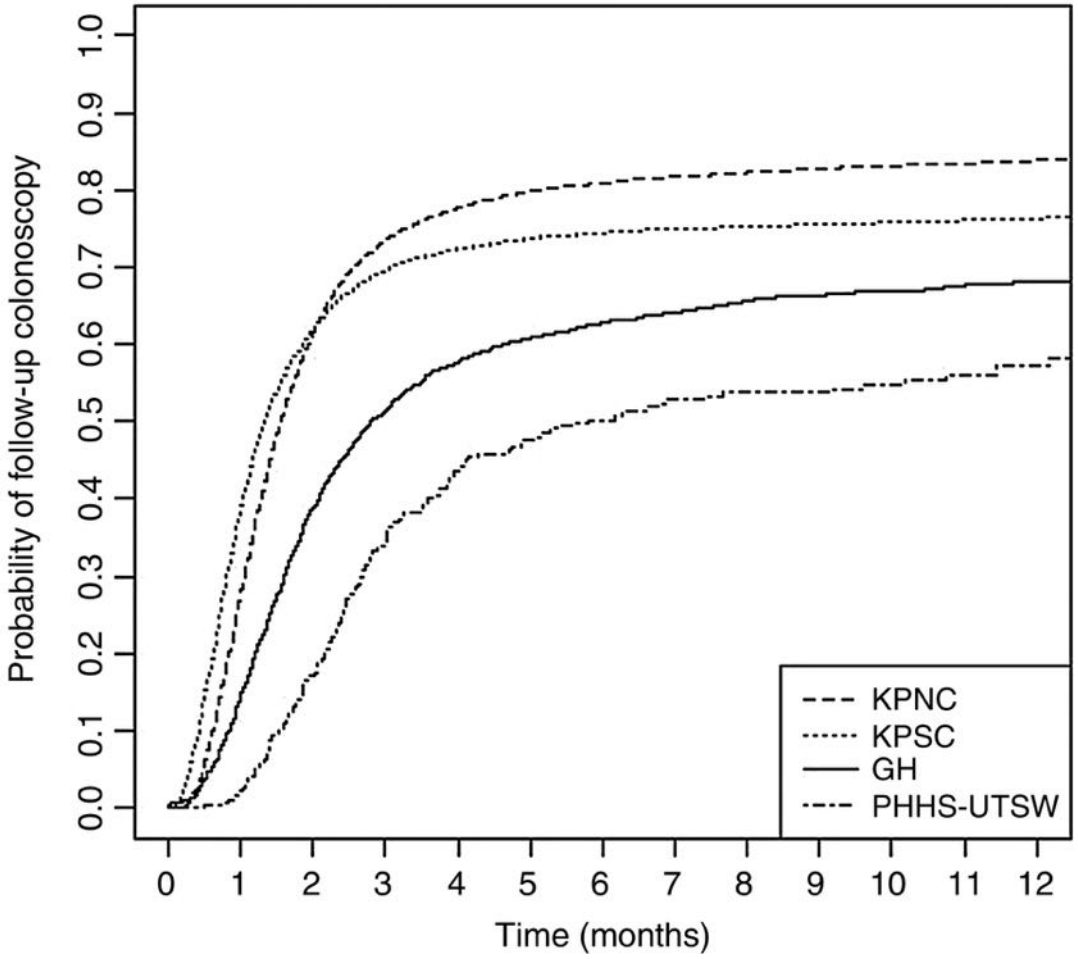


Positive screening stool test

Diagnostic colonoscopy

KAISER PERMANENTE®

Follow-up colonoscopy after positive stool test



Chubak et al. Cancer Epidemiol Biomarkers Prev 2016;25:344-350

KAISER PERMANENTE®

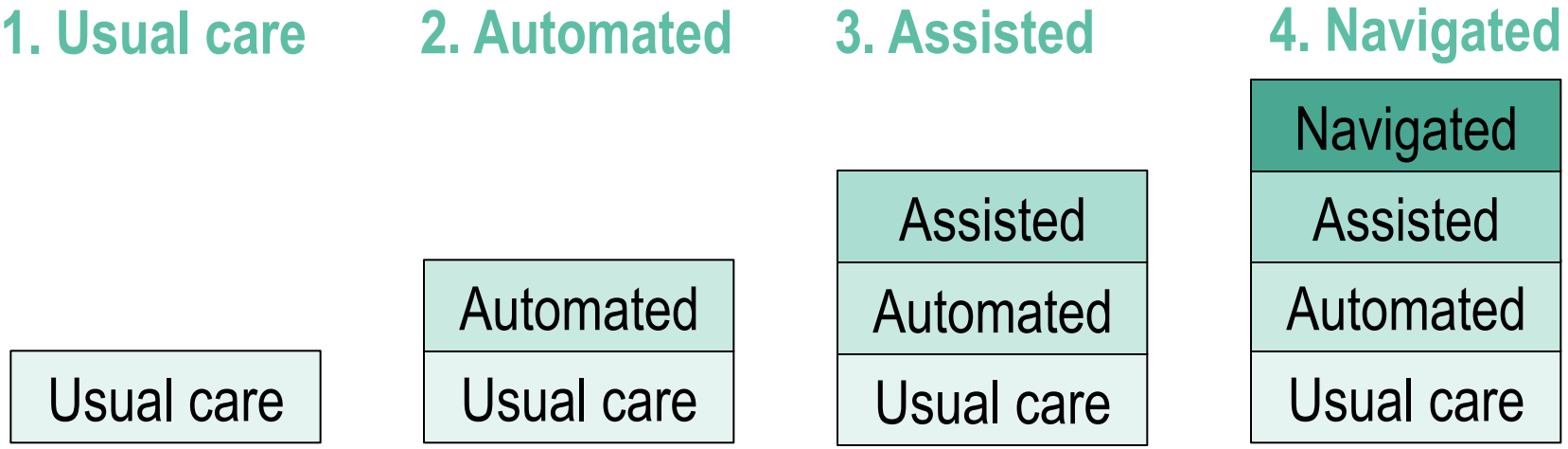
Example 3: Platform for interventions
(What we *can* do)

Annals of Internal Medicine | ORIGINAL RESEARCH

An Automated Intervention With Stepped Increases in Support to Increase Uptake of Colorectal Cancer Screening

A Randomized Trial

Beverly B. Green, MD, MPH; Ching-Yun Wang, PhD; Melissa L. Anderson, MS; Jessica Chubak, PhD, MBHL; Richard T. Meenan, PhD; Sally W. Vernon, PhD; and Sharon Fuller, BA



KAISER PERMANENTE®

Effect of intervention on CRC testing

	% Current for CRC testing over 2 years (95% CI)
Usual care	26.3 (23.4-29.2)
Automated	50.8 (47.3-54.4)
Assisted	57.5 (54.5-60.6)
Navigated	64.7 (62.5-67.0)

Ann Intern Med. 2013;158(5_Part_1):301-311. doi: 10.7326/0003-4819-158-5-201303050-00002

Part 2

PRACTICAL CONSIDERATIONS

KAISER PERMANENTE®

Advantages of claims- and EHR-based research

- Mitigation of selection bias
- Generalizability
- Not subject to recall bias

Health care data

Wha

https://www.mar

KAISER PERMANENTE®

Data sources

Claims	• Procedure codes • Diagnosis codes
Structured clinical data	• Forms • Laboratory results
Unstructured clinical data	• Pathology report text • Colonoscopy reports

Data needed for CRC screening research

- Test indication
- Tests results
- Cancer incidence
- Cancer mortality
- Cancer risk factors

Challenges with using health care data

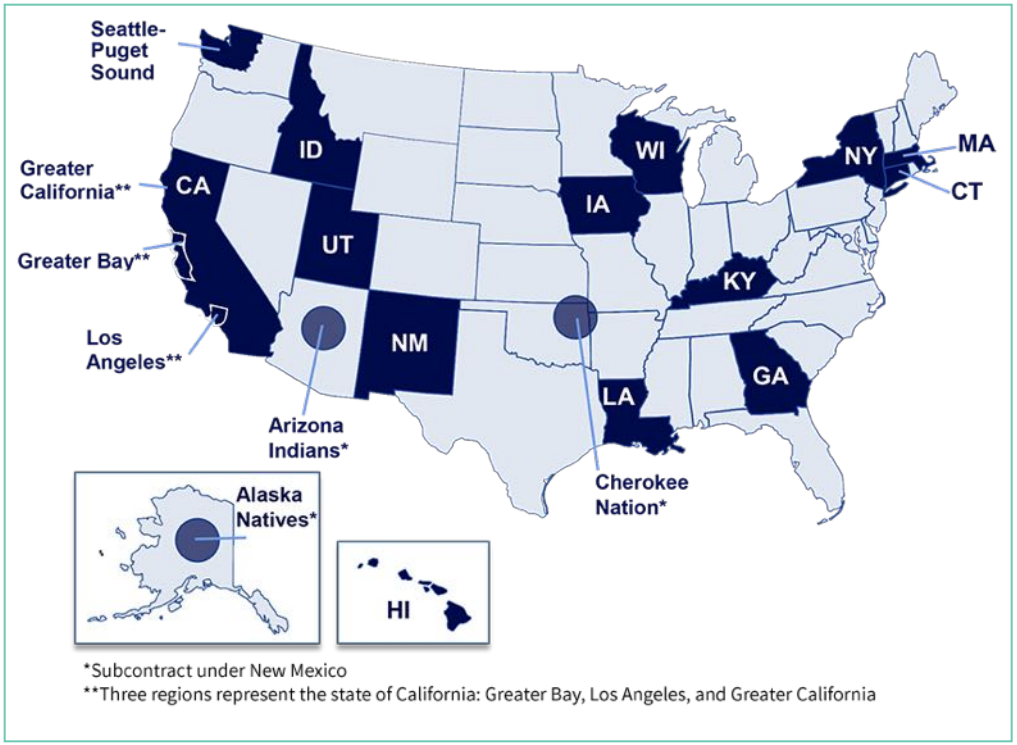
	Claims	Structure clinical	Text
Test indication	No	Sometimes	Yes
Tests results	No	Sometimes	Yes
Cancer incidence	Limited accuracy	No	Yes
Cancer mortality	No	No	Sometimes
Cancer risk factors	Sometimes	Sometimes	Sometimes

Opportunities for getting the necessary data

1. Link data sources
2. Develop and validated EHR-based algorithms
3. Find or create structured data
4. Use unstructured data (i.e., text)

1. Data linkages

KAISER PERMANENTE®



https://seer.cancer.gov/registries/

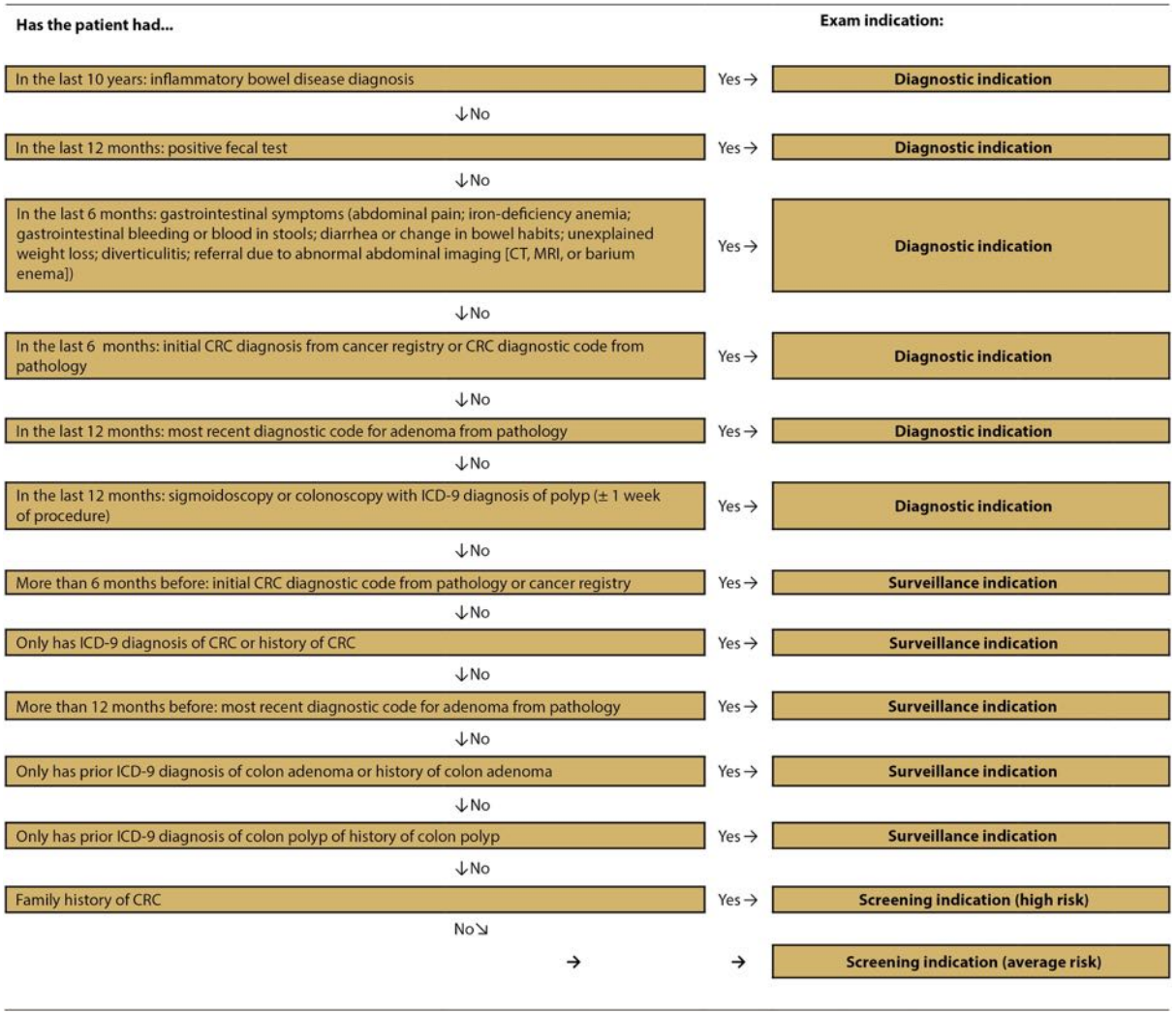
KAISER PERMANENTE®

Caveats with linkages

- Not everyone can be linked
- Requires sharing direct identifiers
- Linkages aren’t always perfect.

2. EHR algorithms for cancer screening research

Example: colonoscopy indication



Diagnostic

Surveillance

Screening

Lee JK, et al. . Gastrointest Endosc. 2015;81:575-82 e4

KAISER PERMANENTE®

Caveats with EHR-based algorithms

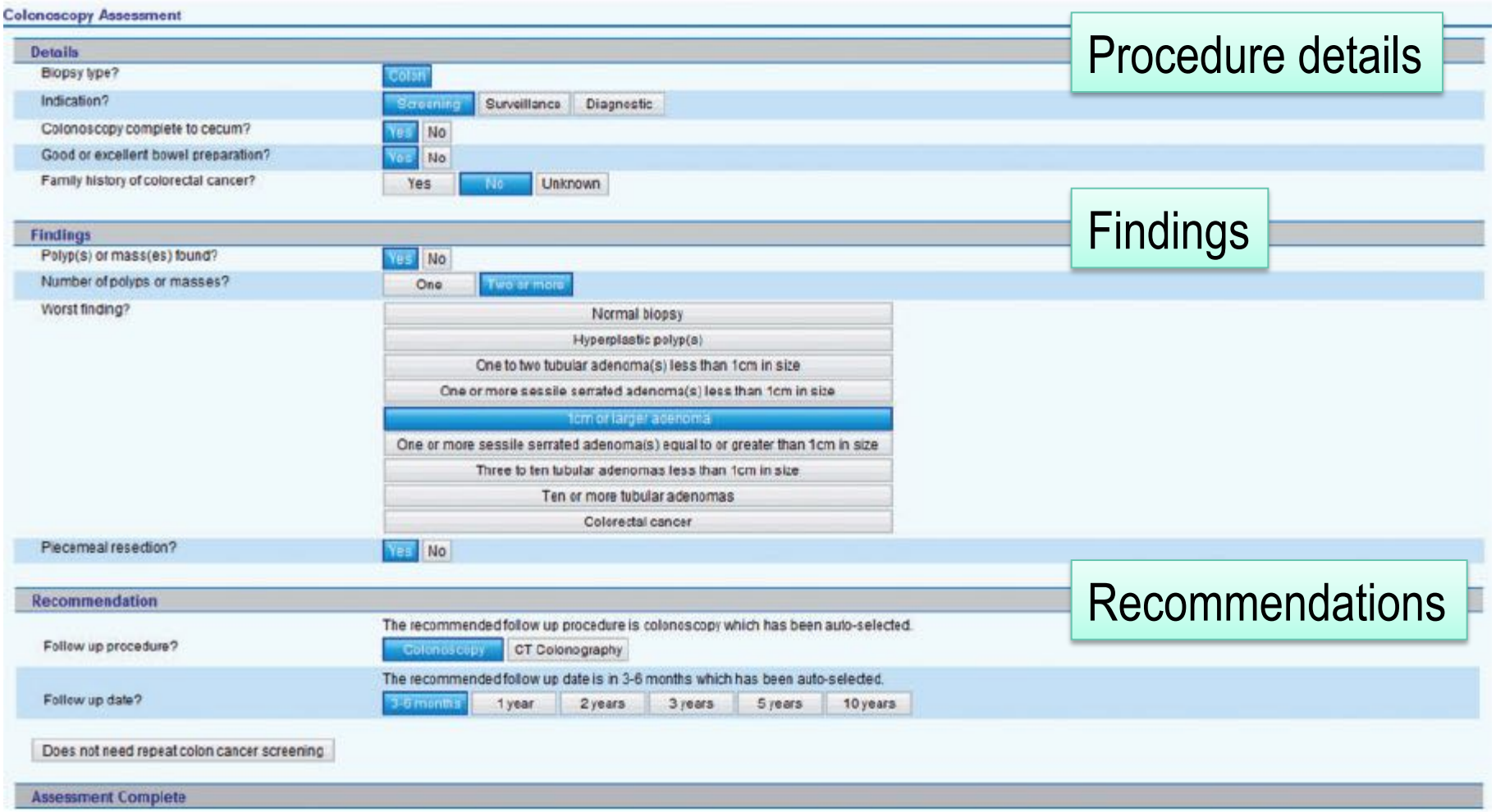
- Misclassification
  - Missing data
  - Coding errors
- Window of data availability
- Different coding practices in different settings
- Switch from ICD-9 to ICD-10

Find or create structured data

- Leverage reporting requirements
  - Mammography Quality Standards Act
  - CMS-approved lung cancer screening registries
- Partner with care providers and delivery systems
  - Develop reporting systems

Example: Colonoscopy Reporting System

Figure 1: Parkland-UT Southwestern CoRS screen shot.



Skinner CS, et al. J Am Med Inform Assoc. 2016;23:402-6

KAISER PERMANENTE®

Use unstructured data

Chart abstraction

Natural language processing

Example: NLP for high grade dysplasia

Algorithm

1.Look for key words (e.g., dysplasia, dysplastic)
2.Exclude if preceded by negation key words (e.g., no evidence)
3.Include if preceded by “high grade” key words

Caveats with text

- Not always available
- Expensive to work with
- Charting practices (and definitions) can be inconsistent

Part 3

SUMMARY AND RECOMMENDATIONS

KAISER PERMANENTE®

Summary

- Healthcare data are critical for improving CRC screening
- Healthcare data offer many advantages
- Data are not always “research ready”

Recommendations

1. Leverage the advantages of health care data
2. Know your source data and its limitations
3. Consider different data collection approaches
4. Validate your approaches
5. Don’t go it alone: find networks and consortia

Thank you

jessica.chubak@kp.org

KAISER PERMANENTE®