

ACS Cancer Conference 2025

March 12-14 | Phoenix, AZ

Ensuring Quality Data Capture and Entry

Matthew HG Katz MD, FACS, FASCO

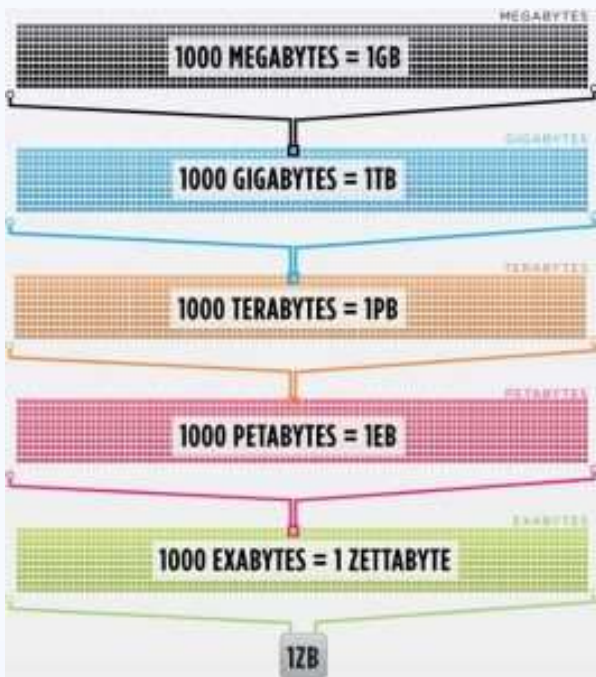
Professor and Chair

Department of Surgical Oncology

The University of Texas MD Anderson Cancer Center

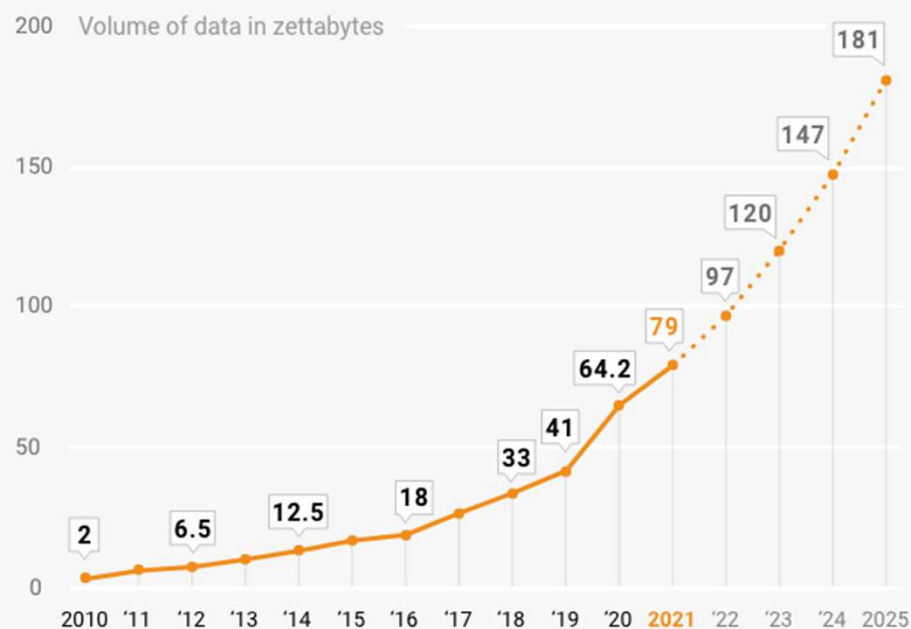


181 zettabytes and rising (fast)



**403
million
terrabytes
of data
generated
daily**

Volume of data created, captured, copied and consumed worldwide



© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Statista.com

Trial design complexity is driving data growth

- Endpoints for phase II/III trials increased 27% since 2009, to 20
- Procedures/pt in phase II/III trials increased 44% since 2009, to 263
- Phase III trials generate 3x data since 2009: 3.6 million data points/trial
- Data more diverse: tissue, labs, PROMs, remote monitoring devices etc.

Why does data quality matter? Ethics

- Valid data essential for sponsors and researchers to make informed decisions re: patient care
- Data must be of sufficient quality and integrity to support endpoint analyses
- Poor data results in a waste of effort and resources
- Poor data may put subjects and patients at risk for harm
- **Subjects are donating their time and bodies to answer our research questions**



© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

GCP guidance for data management

“The quality and amount of the information generated in a clinical trial should be sufficient to address trial objectives, provide confidence in the trial’s results and support good decision making. “



INTERNATIONAL COUNCIL FOR HARMONISATION OF TECHNICAL
REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED GUIDELINE
GUIDELINE FOR GOOD CLINICAL PRACTICE
E6(R3)

Final version
Adopted on 06 January 2025

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

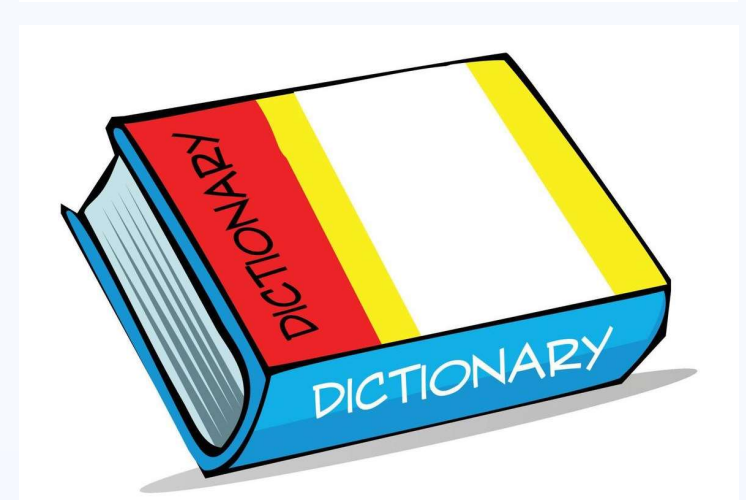
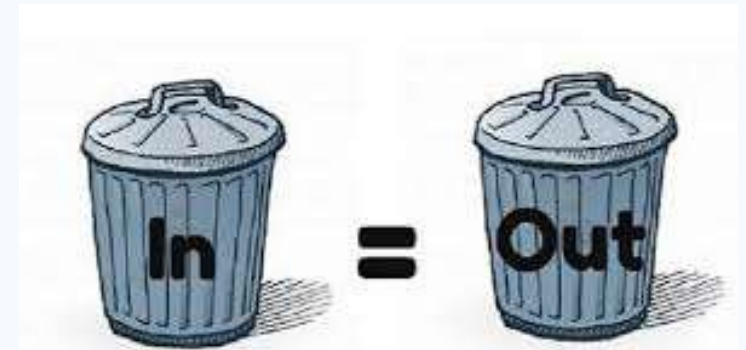
Develop a data management plan (DMP)

- Outlines how data will be collected, managed, assessed, stored and preserved
- Includes:
 - Specification of the data to be collected
 - Case Report Forms
 - Best practices for collecting source docs, extracting data to CRFs, maintaining confidentiality, data storage
 - Processes for QA and QC

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Data Standardization

- Clear data standards and definitions
- Standardized terminology and coding systems
- Surgical data are a function of technical process and context - begin with the surgeon and pathologist

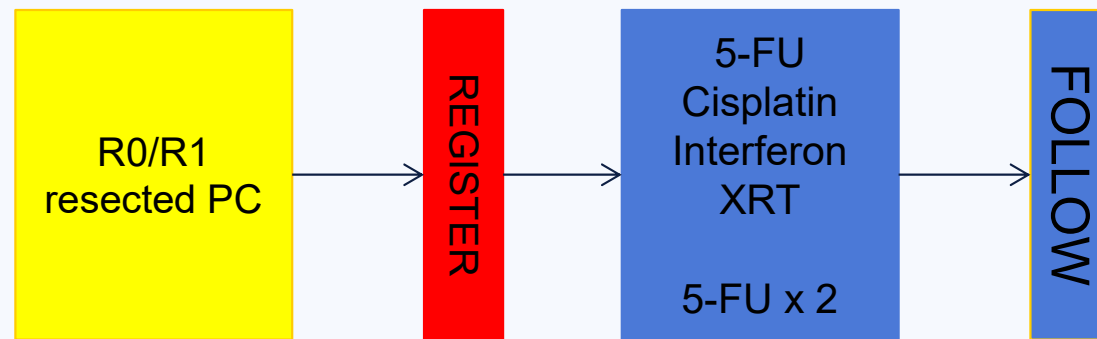


Element name, data type, responses, source, definition, relationships, security

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

ACOSOG Z5031

Cooperative group trial of a postoperative RT regimen for pancreas cancer
Regimen previously evaluated at a single institution – results incredible!



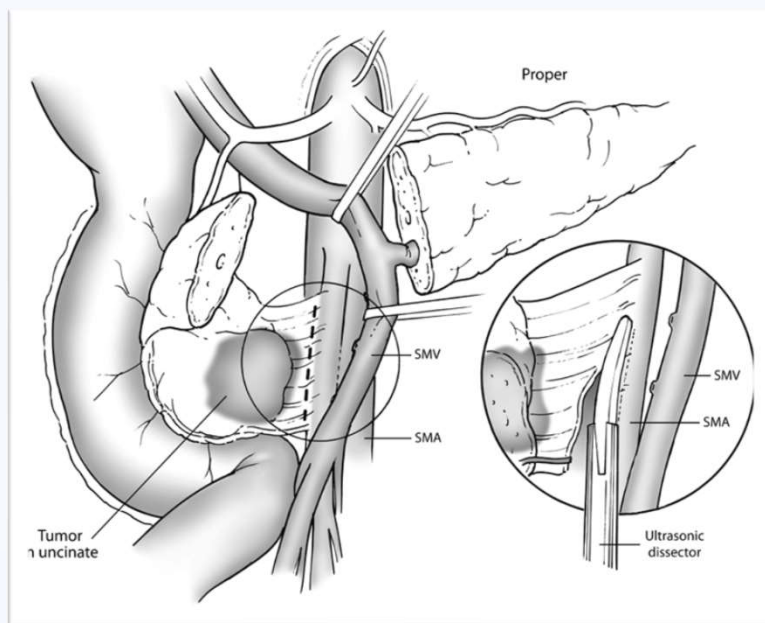
Protocol similar to previous single-institution study: all patients had pancreas cancer, ECOG PS 0-1, *reportedly* undergone R0/R1 resection

**Local recurrence rate following treatment much higher than expected:
46% of patients recurred within the surgical field**

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Standardization of Surgical and Pathologic Variables is Needed in Multicenter Trials of Adjuvant Therapy for Pancreatic Cancer: Results from the ACOSOG Z5031 Trial

Matthew H. G. Katz, MD¹, Nipun B. Merchant, MD², Steven Brower, MD³, Megan Branda, MD⁴, Mitchell C. Posner, MD⁵, L. William Traverso, MD⁶, Ross A. Abrams, MD⁷, Vincent J. Picozzi, MD⁸, Peter W. T. Pisters, MD¹ and The American College of Surgeons Oncology Group



Proper SMA dissection: 25%

- Specimen inked: 66%
- Nodes evaluated: 1 – 49
- TNM stage reported: 49%
- Path met CAP guidelines: 34%

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Margins evaluated prior to enrollment in ACOSOG Z5031

Result	No. of patients (%)			
	Panc	CBD	SMA	AJCC*†
Evaluated, n (%)	79 (100)	74 (94)	37 (47)	36 (46)
Positive, n (%)	12 (15)	2 (3)	14 (38)	16 (44)
Negative, n (%)	67 (85)	72 (97)	23 (62)	20 (56)

* Cases in which all three margins—SMA, CBD, Panc—were evaluated as per AJCC guidelines

† Positive indicates at least one margin of the three was positive; negative indicates all three negative

Although only 25% of patients enrolled in Z5031 were reported to have undergone R1 resection, at least one margin was positive in 44% of the 36 patients in whom evaluation of all three AJCC margins was documented

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Design effective data collection tools (CRFs)

- Case Report Forms should be:
 - Accurate
 - Complete
 - User-friendly
 - Clear
 - Easy to complete
 - Completed in a timely manner
- Use controlled vocabularies/choice lists
- Should meet the study's objectives and requirements
- Historically, paper

Alliance for Clinical Trials in Oncology
Protocol Number: A021501

Patient ID: _____
Institution (Inst. Number): _____

Radiographic Imaging Assessment: Baseline Central Review

CT/MRI RESULTS

Was a CT or MRI of the chest performed? ☐ Yes ☐ No
(If yes), date of CT or MRI of the chest (dd MMM yyyy) ____-____-____

Was a CT or MRI of the abdomen performed? ☐ Yes ☐ No
(If yes), date of CT or MRI of the abdomen (dd MMM yyyy) ____-____-____

(If yes), primary tumor identified (check one) ☐ Yes ☐ No ☐ Not Evaluable
(If yes), primary tumor size ____ mm x ____ mm

CRITERIA FOR BORDERLINE RESECTABLE DISEASE
(The following must be true)

There is no interface between the tumor and the aorta (check one) ☐ Yes ☐ No

(Also, at least one of the following has to be "yes".)

An interface is present between the primary tumor and the superior mesenteric vein or portal vein and measures $\geq 180^\circ$ of the circumference of the vessel wall (check one) ☐ Yes ☐ No ☐ Not Evaluated

Short-segment occlusion of the SMV-PV is present with normal vein above and below the level of obstruction that is amenable to resection and venous reconstruction (check one) ☐ Yes ☐ No ☐ Not Evaluated

Short segment interface (of any degree) is present between tumor and hepatic artery with normal artery proximal and distal to the interface that is amenable to resection and reconstruction (check one) ☐ Yes ☐ No ☐ Not Evaluated

An interface is present between the tumor and superior mesenteric artery or celiac axis measuring $<180^\circ$ of the circumference of the vessel wall (check one) ☐ Yes ☐ No ☐ Not Evaluated

RADIOGRAPHIC STAGE

Radiographic stage ☐ Resectable
☐ Borderline resectable
☐ Locally advanced
☐ Metastatic
☐ Not evaluable

BASELINE CENTRAL REVIEW RESULT

Patient confirmed eligible by central radiographic review: (check one) ☐ Yes ☐ No

Comments

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

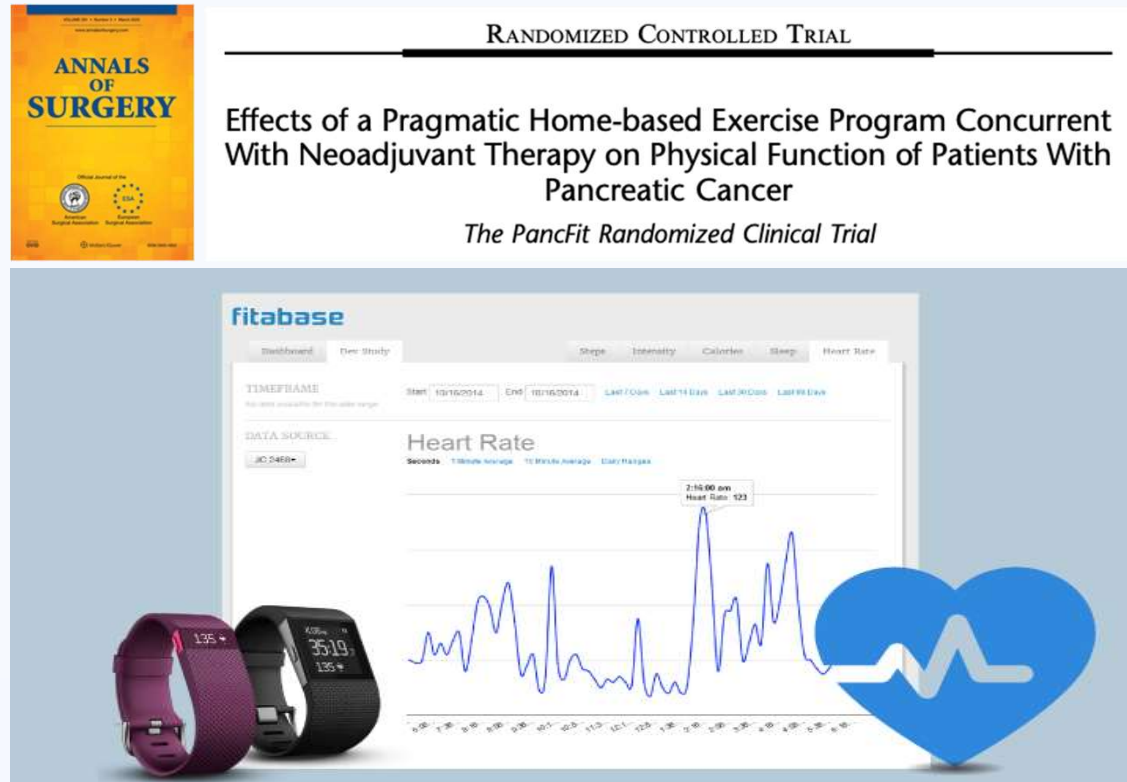
Electronic Data Capture Systems (eCRFs)

- Streamline data collection, storage and management
- Offer real-time data entry
- Improve accuracy and security
- Reduces manual data entry/transfer errors and facilitate review/analysis
- Include data and metadata (signatures, times, audit trail)



© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Automated Data Entry

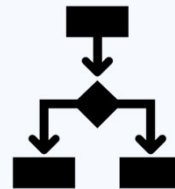


© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Synoptic Operative Reporting



**Standardized data elements
organized as a structured
checklist
or template**



**Each data element's value is "filled
in" using a pre-specified format to
ensure interoperability of
information**

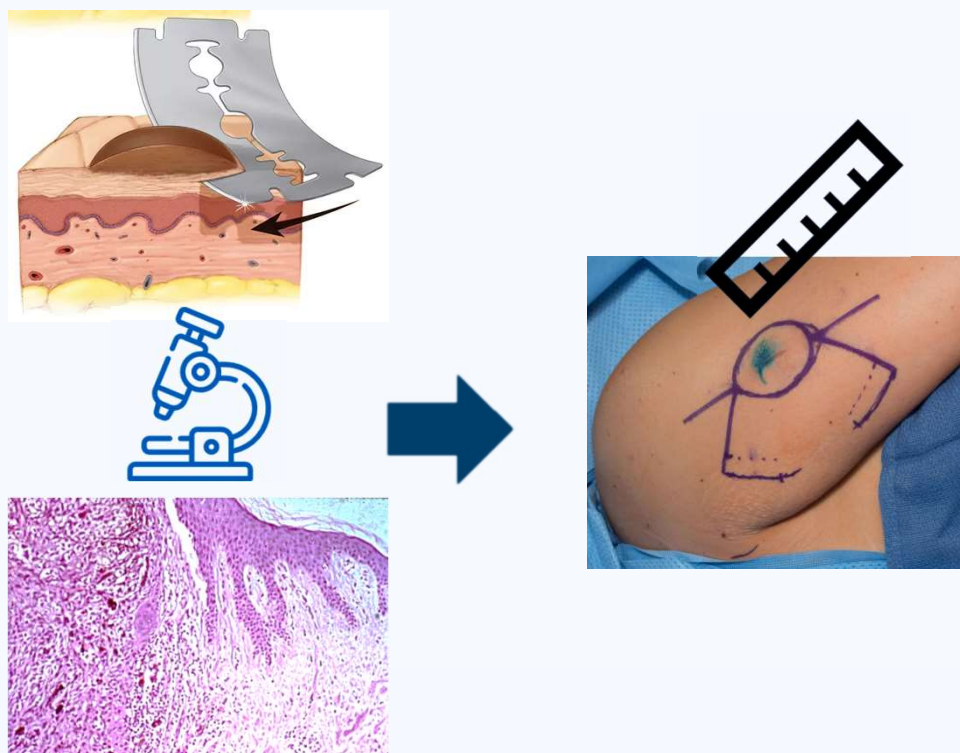
- **The information being sought is standardized**
- **The options for each variable are constrained to a pre-defined set of responses**



**Synoptic reports allow
information to be easily
collected, stored, and
retrieved**

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

CoC Standard 5.5 - Wide Local Excision Melanoma



Element	Response Options
Operation performed with curative intent	Yes; No.
Original Breslow thickness of the lesion	Melanoma in situ (MIS); __ mm (to the tenth of a millimeter).
Clinical margin width (measured from the edge of the lesion or the prior excision scar)	0.5 cm; 1 cm; 2 cm; Other: __ cm due to cosmetic/anatomic concerns; Other (with explanation).
Depth of excision	Full-thickness skin/ subcutaneous tissue down to fascia (melanoma); Only skin and superficial subcutaneous fat (melanoma in situ); Other (with explanation).

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Narrative Reporting vs. Synoptic Reporting

Narrative reporting...

- May be constructed using pre-determined data fields and pre-determined responses
- Constructed by dictation, free text, smarttext, etc.
- May use standardized terminology
- Presented in a **prose** format
- Prone to **omission** of necessary data and **inconsistencies** in language and formatting
- May allow for discrete data capture

Synoptic reporting...

- **Always** constructed using pre-determined data fields and pre-determined responses
- Typically created using a **tool**
- **Always** uses standardized terminology
- Presented in **checklist** format
- **Always** allows for discrete data capture
 - Information is formatted so it can be collected, stored, and is easily retrievable for data repositories
 - Can automatically populate data from the EHR

A note may (ideally?) be a combination of the two!

Quality Assurance

- Planned activity to ensure quality requirements are fulfilled
- Manual and automated procedures
- Should be specified in the DMP
- Data collection
 - Multiple measurements
 - Standardized methods for capturing observations
 - Computer-assisted software
- Data entry
 - Validation rules or input masks in data entry software
 - Data entry screens
 - Detailed labeling



© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Quality Control

- Review of data to verify accurately, validity, completeness
 - Verification of random samples
 - Histograms to detect anomalous values
 - Peer review
- Performed by research staff
- Continuous, in real time, close to data collection



© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Training and knowledge sharing

- Practical aspects of role
- Importance of data integrity
- Collection and management procedures
- Use of software and equipment
- Review and update training materials, training refreshers



© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Responsibility

- PI ultimately responsible for data quality and integrity
- Investigators responsible for collecting and recording data
- Staff responsible for implementing the DMP

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.

Thank you

- *Matthew Katz: mhgkatz@mdanderson.org*

© American College of Surgeons. Content cannot be reproduced or repurposed without written permission of the American College of Surgeons.