

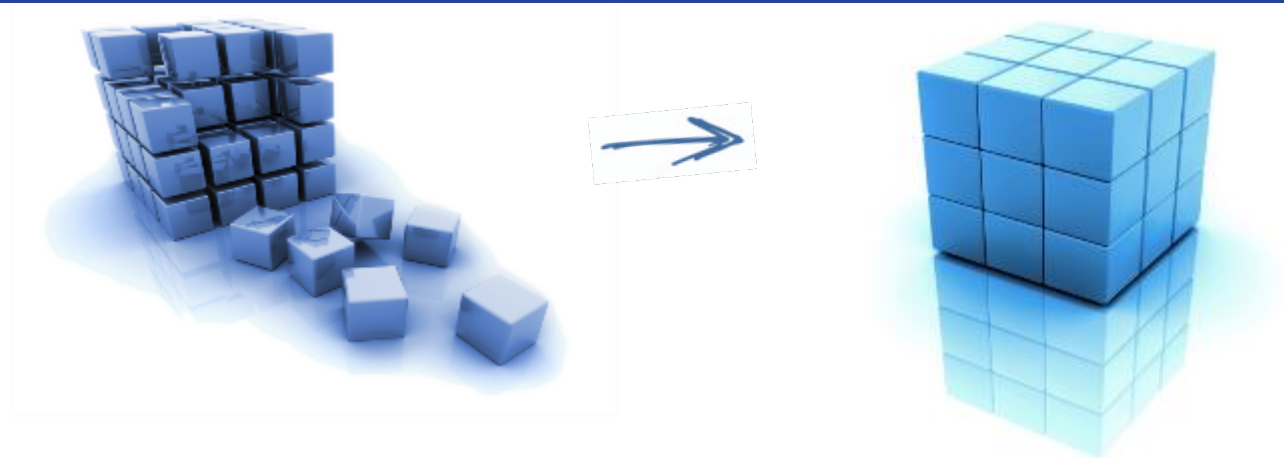
Validation versus Consistency Checking of SEND datasets

Mohit Mathew, Senior Architect

Head of eDataValidator Program (SDTM, SEND, ADaM) and MySEND

Architect of Risk Modeling & Simulation Tool, CTP, FDA

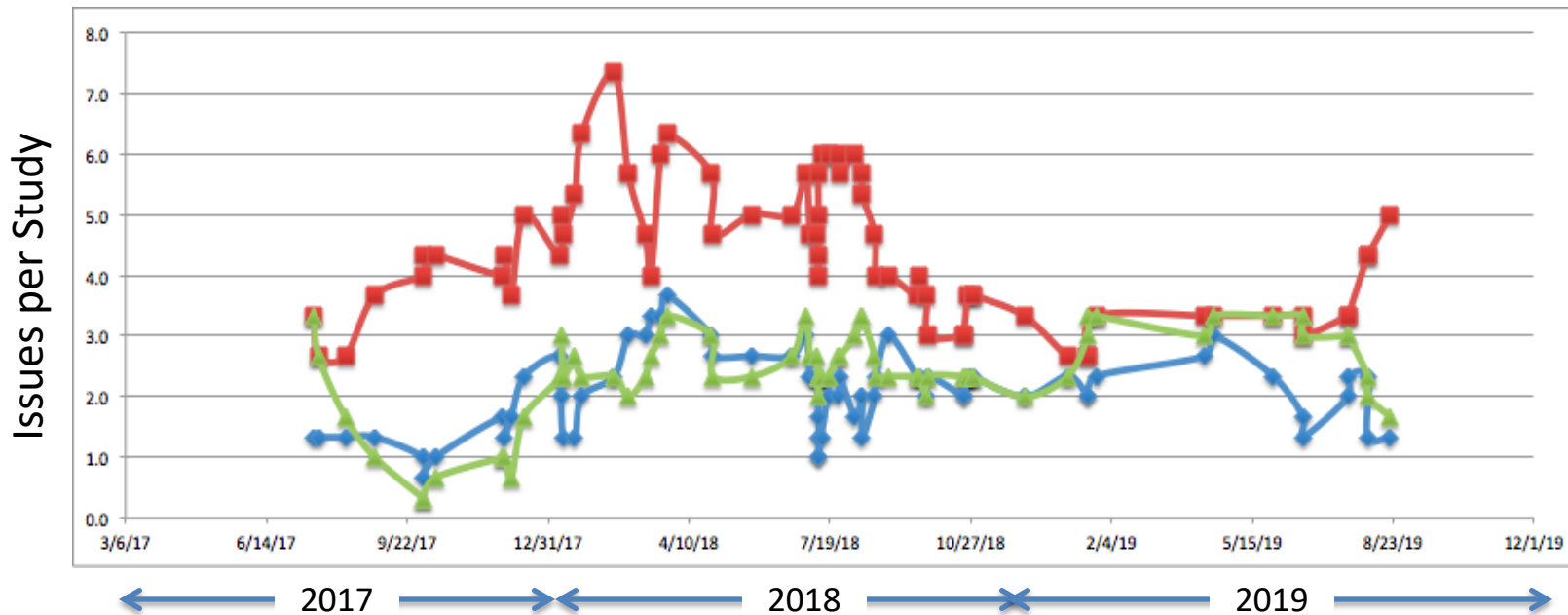
Architect of SEND Data Standardization and Assurance Tools & Processes



www.pointcross.com

How big a problem is it ?

- Out of 69 Studies from SEND-ASSURE checks – NOT ONE STUDY was considered Fit for Reviewers without Remedy
- 100% checking for consistency IS POSSIBLE . IS COST-EFFECTIVE . CAN GENERATE REMEDIAL Corrections
- Biggest issue is lack of consistency with Study Report and Inability to re-generate Study Report Summaries
- Conformance can be checked with an eDataValidator easily – but Consistency cannot be easily verified



- Consistency issue between send and pdf
- Quality and best practice issue
- Conformance to standards

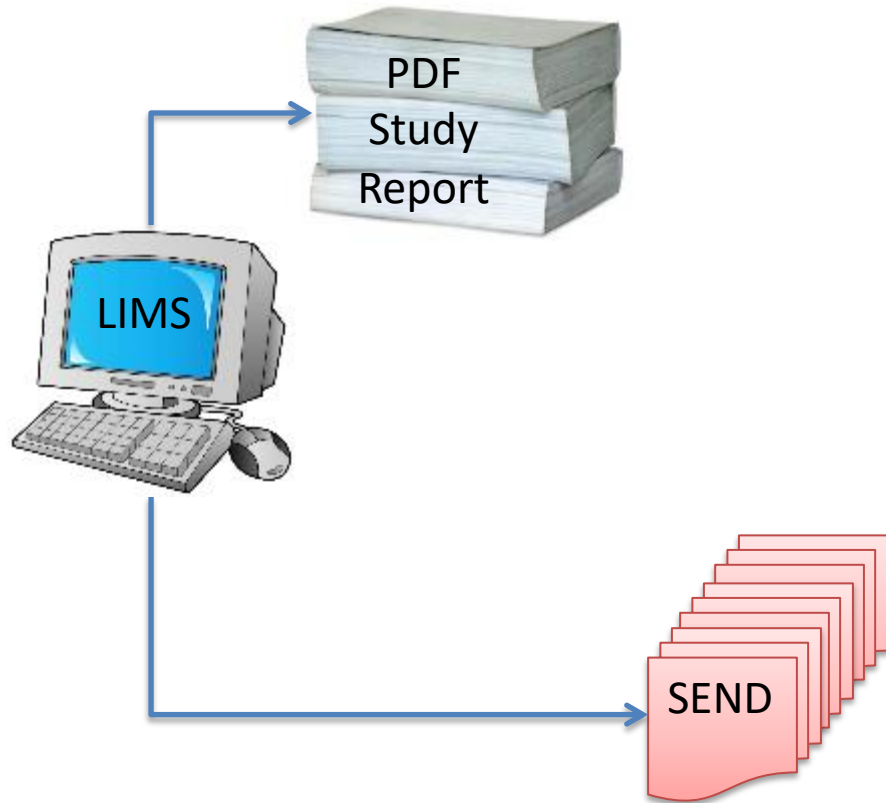
When is a SEND dataset Fit for Review?

- Passes the Technical Rejection Criteria – routine process at eCTD & eData
 - ✓ <https://www.fda.gov/media/100743/download>
- Passes SEND and Define.xml conformance validation using a Validator - Routine
- Passes the Fitness for Review assessment such as the FDA Kick-Start assessment
 - ✓ This is difficult, time consuming and quite expensive for the FDA
 - ✓ Can SEND preparers check for data consistency between SEND and Study Report? If not
 - the SEND dataset is not fit for review
 - ✓ Can SEND preparers ensure that every published summary statistic and incidence count can be generated by the SEND data set? If not – the SEND dataset is not fit for review

SEND Data conformance validation

- Easily done with a software Validator tool
- Validators check conformance to published RULES for SEND and define xml standards. Non-conformant datasets cannot and should NOT be reviewed
- CDISC, FDA and PMDA Rules are published, and transparent providing a clear path to ensure conformance.
- Conformance validation **does not** ensure that SEND dataset is →
 - ✓ Accurate?
 - ✓ Complete?
 - ✓ Consistent with the Data in the PDF report?
 - ✓ Capable of generating at least the same summary results – statistics and incidence counts as in PDF report?
- That is why a 100% Consistency Check is Essential before a SEND dataset is considered fit for review.

Challenges for SEND: 2 Separate Processes

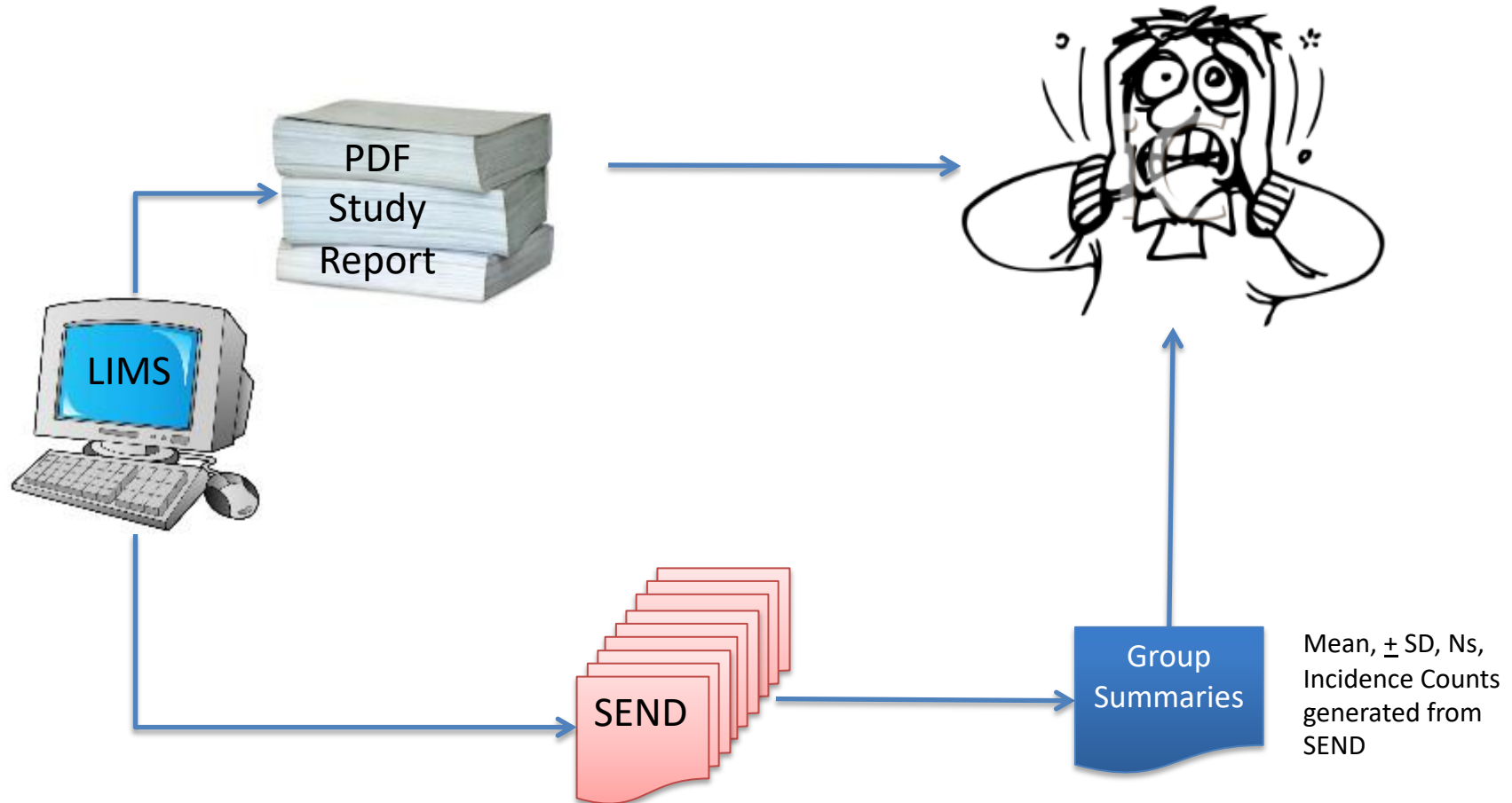


SEND Vs PDF Study Report

- Different Teams
- At Different Times
- Different Study Designs
- With Different Terminology
- Subject & Summary in PDF vs Subject Data alone in SEND
- Study Director does not review SEND datasets
- No ADaM – therefore it is up to the SEND team if they can regenerate the summaries in the PDF – very difficult
- How can FDA reviewers be confident that SEND datasets accurately represent data in PDF reports when there is no way to certify each SEND data set against the Study Report?

Manual Spot Check?

7,000-10,000 sets of summary numbers in a typical study
500,000 to 1M or more for subject level data
How to figure out Grouping and Timing assumptions in PDF?

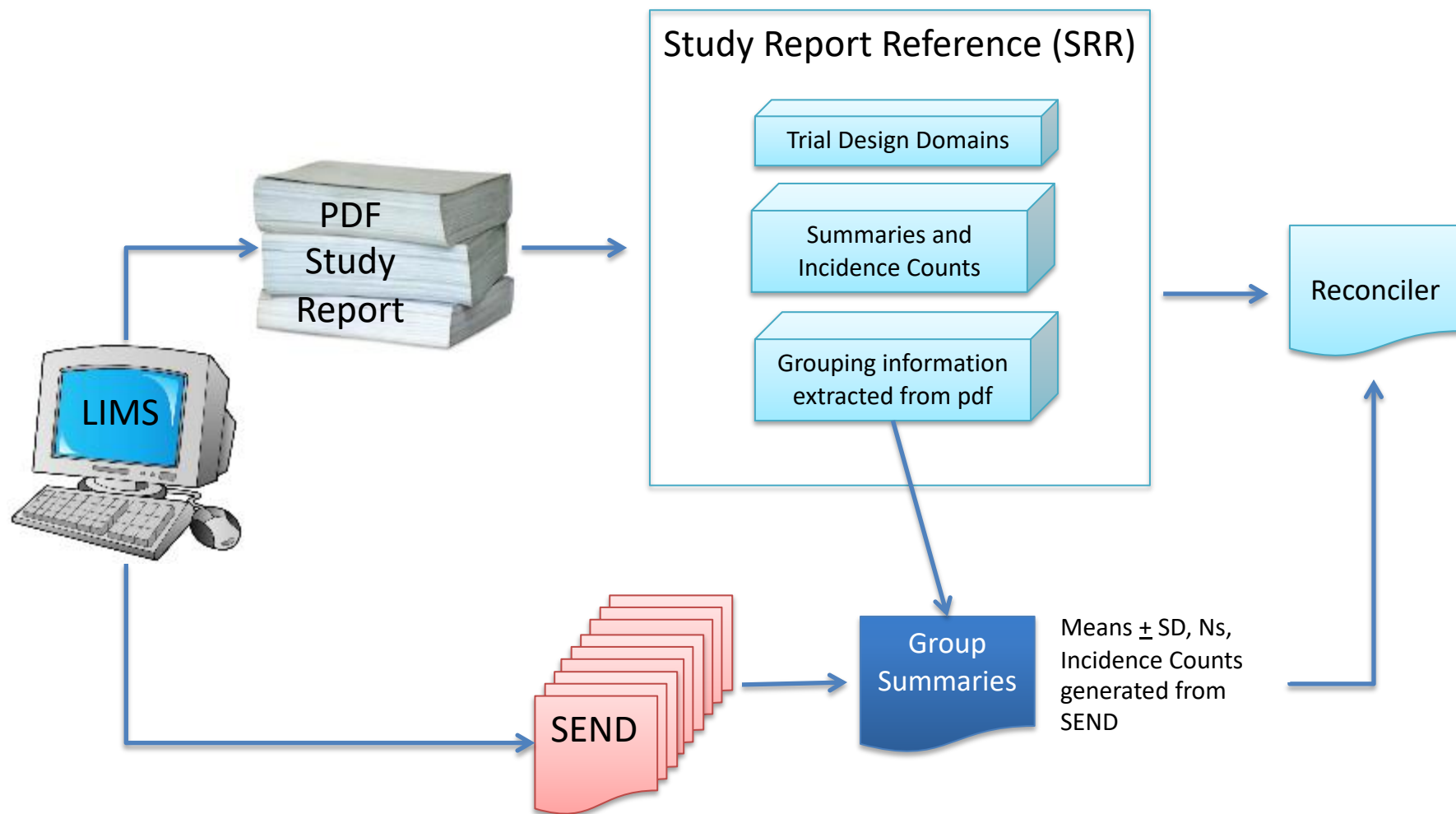


Conformance validation + Spot check is NOT enough.

- Spot checks comparing SEND with PDF are not enough!
 - Out of 10,000 summaries published in the PDF – HOW MANY should be SPOT Checked to get 100% confidence? 90% confidence? 80% confidence?
 - Will a Reviewer make a Regulatory assessment with data that has been spot checked at 90% confidence?
 - We don't believe so.
- FDA reviewers have to be assured that submitted SEND data sets used with their validated tools and scripts are safe to make regulatory decisions
- The proof is in the SEND dataset being capable of re-generating ALL the published summaries in the PDF report
 - ✓ If it can't – then figure out the Root Cause before using the SEND dataset
- 100% comparison of SEND against PDF Reports is possible if the PDF can be represented by digital, columnar, tabulated machine readable files

Solution: Automated, 100% Consistency Check

- Represent Study Report as a Machine Readable File (SRR)
- Automatically Reconcile SRR against SEND – Let a computer do the dumb work



SRR : Generating the Trial Design for reconciliation

- Generate a Trial Summary Using ONLY the Study Report
- Independently generate the Trial Design domains using ONLY the Study Report
 - ✓ PointCross provides a free tool for this called MySEND.
- Use this TE, TA, TX, DM and EX to validate the SEND structure of reporting before checking and reconciling data consistency

✈ Trial Design Editor - PC201904

SEND Study Design Editor: PC201904								
Trial Summary		TD Generation Form	Trial Elements	Trial Arms	Trial Sets	Demographics	Subject Elements	Exposure
+ Add Delete Save Refresh Copy								
☐	TSSEQ *	TSGRPID	TSPARM *	TSPARMCD *	TSVAL *			
☐	1		Study Title	STITLE	13-Week Repeat Dose Toxicity Study on PCDRUG1234 in Rats			
☐	1		Study Design	SDESIGN	PARALLEL			
☐	1		Study Type	SSTYP	REPEAT DOSE TOXICITY			
☐	1		Investigational Therapy or Treatment	TRT	PCDRUG1234			
☐	1		Treatment Vehicle	TRTV	Saline			
☐	1		Route of Administration	ROUTE	ORAL GAVAGE			
☐	1		Dosing Duration	DOSDUR	P13W			
☐	1		Species	SPECIES	RAT			

SRR : Generate a Study Reference Summary (SRS)

Findings

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Section	Group	Sex	Test	Specimen	Unit	Mean	SD	SE	Count	Day	Time Point	Remarks	Page Number	Table Number
2	Summary of Hematology Data	1	m	RBC		M/uL	7.2	0.115		5	3	predose		49	10
3	Summary of Hematology Data	1	m	HGB		g/dL	13.5	0.34		5	3	2H		49	10
4	Summary of Hematology Data	1	m	HCT		%	42.4	1.82		5	3			49	10
5	Summary of Hematology Data	1	m	MCV		fL	65.8	1.94		5	W1			49	10

Disposition

	A	B	C	D	E	F	G	H	I
1	Section	Group	Sex	Animal status	Count	Day	Remarks	Page Number	Table Number
2	Animal Disposition	1	Female	Terminal Kill	3	15		33	1
3	Animal Disposition	1	Female	Recovery Kill	2	15		33	1
4	Animal Disposition	2	Female	Terminal Kill	1	15		33	1
5	Animal Disposition	2	Female	Recovery Kill	4	15		33	1

Observations

	A	B	C	D	E	F	G	H	I	J	K	L
1	Section	Group	Sex	Specimen	Observation	Severity	Count	Examined	Day	Remarks	Page Number	Table Number
2	Summary of Histopathology Observations	1	FEMALES	kidney;	cast; proteinaceous; focal	minimal	1	3	15		75	15
3	Summary of Histopathology Observations	1	FEMALES	kidney;	cast; proteinaceous; focal	mild	1	3	15		75	15
4	Summary of Histopathology Observations	1	FEMALES	kidney;	fibrosis; unilateral	marked	1	3	15		75	15
5	Summary of Histopathology Observations	1	FEMALES	brain;	Within Normal Limits		3	3	15		75	15

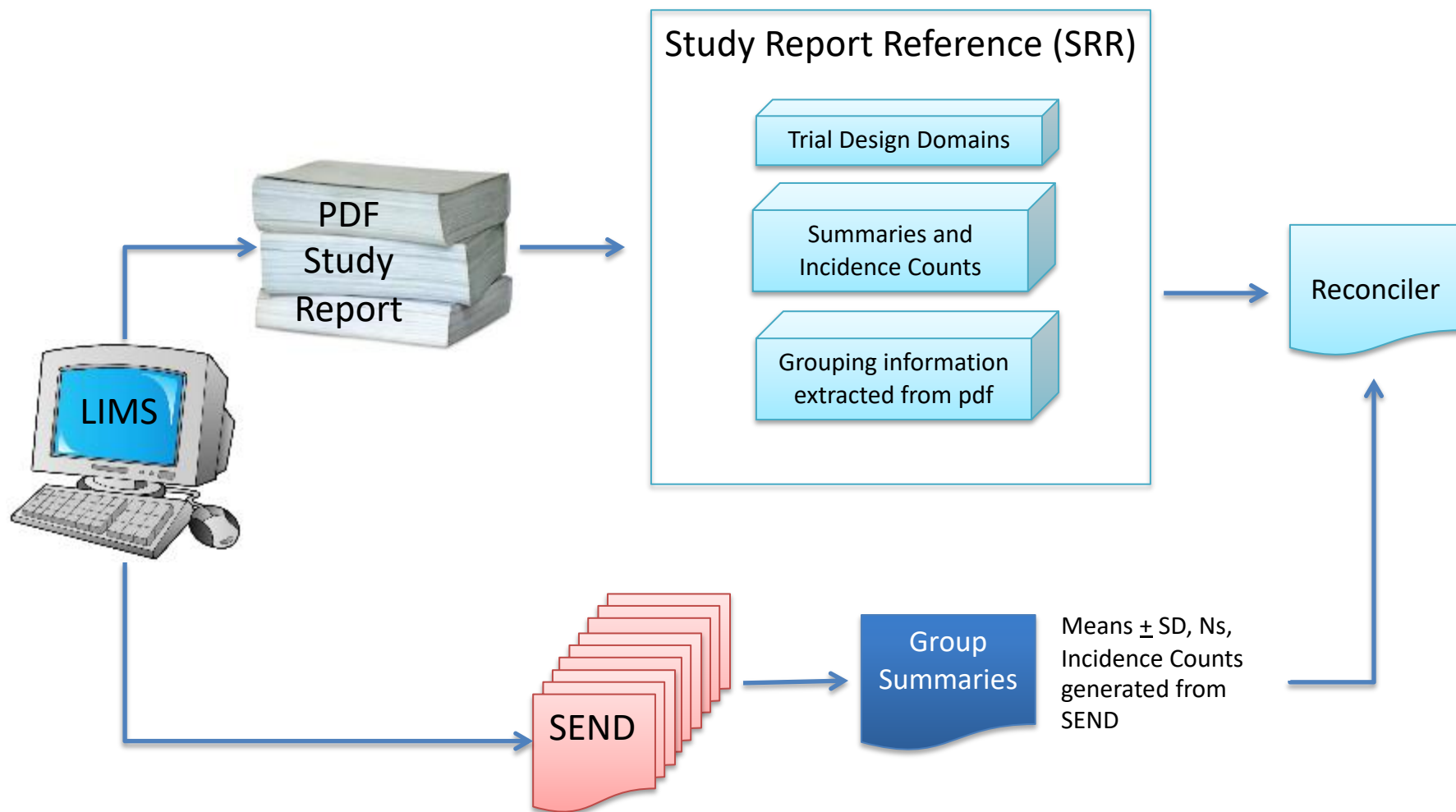
Study Summary

Sponsor's Study Reference No.	PC202020
Study Title	4-Week Repeat-Dose Toxicity Study of PC010 in Rats
Study Type	Repeat Dose Toxicity
Investigational Therapy or Treatment	PC010
Treatment Vehicle	Saline

Means, SDs, Ns for numeric data; observations for clin obs, postmortem findings are extracted from PDF report using OCR, un-pivoted into columnar tabulation. Metadata from PDF tables are included.

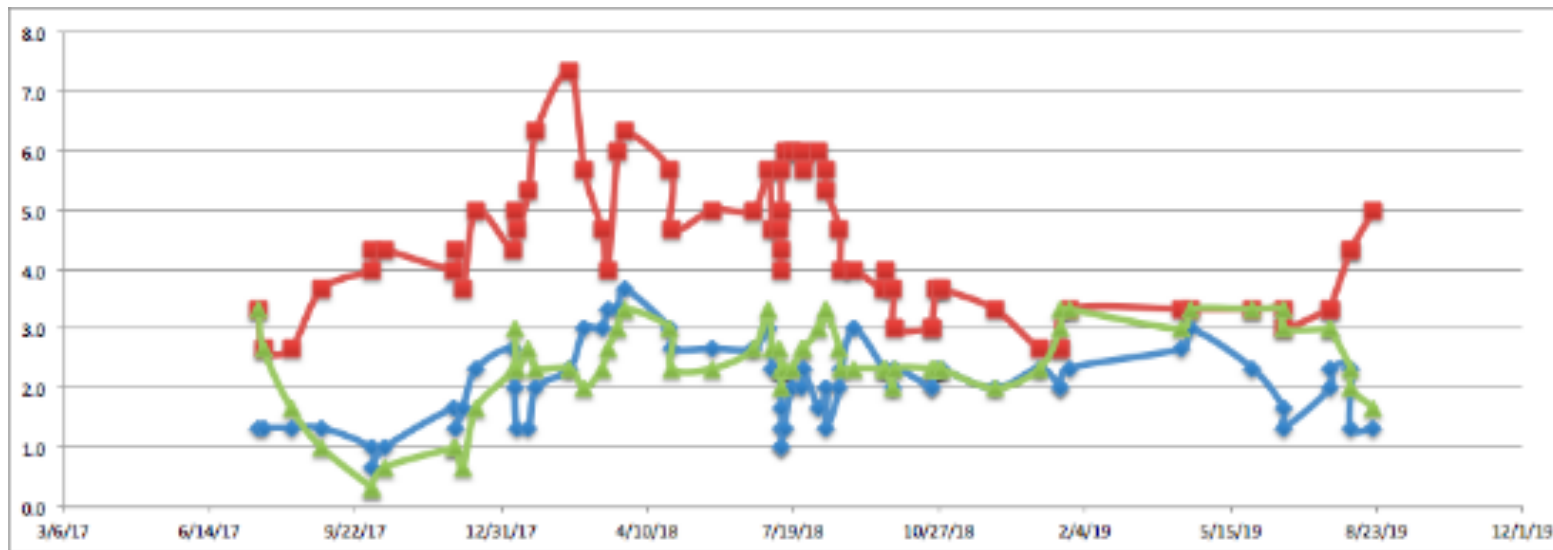
Solution: Automated, 100% Consistency Check

- Represent Study Report as a Machine Readable File (SRR)
- Automatically Reconcile SRR against SEND – Let a computer do the dumb work



Conclusion

- Validators can easily check for conformance
- We see consistency issues in 100% of SEND Datasets where a 100% consistency check was done
- It is cost and time effective to do a 100% consistency check of SEND and PDF study report with automated process and tools
- Consistency issues can be remedied only if they are found.



Thank you, and acknowledgments

Dr. Laura Kaufman DABT,

Dr. Karen Porter, Chief Toxicologist

Karuna Polavarapu, SEND specialist, Bioinformatics

Contact us at mohit@pointcross.com, regulatory@pointcross.com

Tel: +1-844-382-7257

Scan QR Code to get a copy

