



**ASOP 23:
Data Quality**

Annual Meeting
November 2019



ASOP 23: Data Quality

- Considered to apply to all assignments in all areas
- First created in 1993 in response to a case of a reserve opinion based on unaudited data – raised questions regarding actuaries’ responsibilities for:
 - Selecting among alternative data sources
 - Relying on data supplied by others
 - Disclosing deficiencies in the data
- Current version is the 3rd, effective April 2017
 - Clarified to explicitly state that it applies even when preparing data to be used by other actuaries




Key Definitions

What is **data**, anyway?

✓ Numerical, census, or classification info	✗ General or qualitative information
✓ Derived mathematically from other data	✗ Assumptions from other data

Appropriate – has always been defined as

- Suitable, i.e. related for the intended purpose
- Relevant to what is being analyzed



Review ≠ Audit

- What everyone knows about ASOP 23:
You are not required to audit the data.
- But what exactly is the difference?

Review:	Audit:
examination of the <i>obvious</i> characteristics to determine if data appear <i>reasonable</i> and <i>consistent</i>	formal and systematic examination for testing <i>accuracy</i> and <i>completeness</i>
- No need to compile more data solely for the review



The Review

- First and foremost, the actuary should perform a review, i.e. make a reasonable effort to:
 - Determine the definition of each data element used !
 - Identify questionable values or inconsistencies
- And should consider:
 - Comparing to prior data
 - Taking steps to improve the data
- But the actuary is allowed to deem a review not necessary or not practical → disclose why



Considerations for selecting data

- Still aiming for appropriateness of the data, given the *scope and intended use* of the analysis
- Considerations:
 - Appropriate, including sufficiently current
 - Reasonable and internally consistent
 - Consistent with (known) external information
 - The degree to which the data are **sufficient**
 - Known significant limitations
 - Availability / practicality of alternative data



So what if...
your data has some issues?

- ASOP 23 does not preclude the use of imperfect data
- Use your professional judgment whether...
 - The data is good enough or you can correct it
 - You can make judgmental adjustments to the data
 - You can make judgmental adjustments to the results
- But if the data is so inadequate it cannot be used, find different data or decline assignment



So what if...
someone else supplied your data?

- You may rely on others for data, but you're still subject to the ASOP requirements for:
 - Review
 - Judgment call about using the data
- Should disclose reliance on others



Disclosure, disclosure, disclosure...

ASOP 23 has the usual long list of required disclosures:

- Data sources and reliances
- That you reviewed the data, or why not
- Data corrections or judgmental adjustments
- Remaining (material) concerns about data quality
- Implications such as highly uncertain or biased results, and any resulting limitations on the use of the analysis




Kahoot!

Instructions:

- On your phone or tablet, go to **kahoot.it**
- Enter the game pin
- Enter your nickname

Case Study


- You have recently transferred from your company's reserving unit into pricing.
- When you were in reserving, you were comfortable with the various procedures to reconcile the reserving data to the financials, but you did not perform any reconciliations yourself.
- As part of your reserve study, you discussed the occasional outlier with the claims department.
- Now you notice that some of the paid loss data fields pulled in pricing are not the same as the ones you used in reserving (e.g. PDLSCQNC vs. PDLSCQNX) and you realize you don't know what the difference is.
- The pricing department also uses some industry benchmark development patterns from a third party provider.



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?

- Are you really responsible for knowing the definitions of all the data fields?
 - What if the data contained in the fields are identical? Similar? Not even close? Does it matter?
- Does it matter whether you pulled the data yourself or someone else provided it for your analysis?
- Under what circumstances do you NOT have to worry about data quality?
- If you identify an issue with the data, what can you do?
- What if you never get an answer to your question about the paid loss definitions?
- What—if anything—should you have done differently in reserving?
- To what extent are you responsible for understanding the data or definitions underlying the external development patterns?



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?

- Are you really responsible for knowing the definitions of all the data fields?
- **What if the data contained in the fields are identical? Similar? Not even close? Does it matter?**
- Does it matter whether you pulled the data yourself or someone else provided it for your analysis?
- Under what circumstances do you NOT have to worry about data quality?
- If you identify an issue with the data, what can you do?
- What if you never get an answer to your question about the paid loss definitions?
- What – if anything – should you have done differently in reserving?
- To what extent are you responsible for understanding the data or definitions underlying the external development patterns?



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?

- Are you really responsible for knowing the definitions of all the data fields?
- What if the data contained in the fields are identical? Similar? Not even close? Does it matter?
- **Does it matter whether you pulled the data yourself or someone else provided it for your analysis?**
- Under what circumstances do you NOT have to worry about data quality?
- If you identify an issue with the data, what can you do?
- What if you never get an answer to your question about the paid loss definitions?
- What – if anything – should you have done differently in reserving?
- To what extent are you responsible for understanding the data or definitions underlying the external development patterns?



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?


- Are you really responsible for knowing the definitions of all the data fields?
- What if the data contained in the fields are identical? Similar? Not even close? Does it matter?
- Does it matter whether you pulled the data yourself or someone else provided it for your analysis?
- **Under what circumstances do you NOT have to worry about data quality?**
- If you identify an issue with the data, what can you do?
- What if you never get an answer to your question about the paid loss definitions?
- What – if anything – should you have done differently in reserving?
- To what extent are you responsible for understanding the data or definitions underlying the external development patterns?



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?


- Are you really responsible for knowing the definitions of all the data fields?
- What if the data contained in the fields are identical? Similar? Not even close? Does it matter?
- Does it matter whether you pulled the data yourself or someone else provided it for your analysis?
- Under what circumstances do you NOT have to worry about data quality?
- **If you identify an issue with the data, what can you do?**
- What if you never get an answer to your question about the paid loss definitions?
- What – if anything – should you have done differently in reserving?
- To what extent are you responsible for understanding the data or definitions underlying the external development patterns?



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?


- Are you really responsible for knowing the definitions of all the data fields?
- What if the data contained in the fields are identical? Similar? Not even close? Does it matter?
- Does it matter whether you pulled the data yourself or someone else provided it for your analysis?
- Under what circumstances do you NOT have to worry about data quality?
- If you identify an issue with the data, what can you do?
- **What if you never get an answer to your question about the paid loss definitions?**
- What – if anything – should you have done differently in reserving?
- To what extent are you responsible for understanding the data or definitions underlying the external development patterns?



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?

- Are you really responsible for knowing the definitions of all the data fields?
- What if the data contained in the fields are identical? Similar? Not even close? Does it matter?
- Does it matter whether you pulled the data yourself or someone else provided it for your analysis?
- Under what circumstances do you NOT have to worry about data quality?
- If you identify an issue with the data, what can you do?
- What if you never get an answer to your question about the paid loss definitions?
- **What – if anything – should you have done differently in reserving?**
- To what extent are you responsible for understanding the data or definitions underlying the external development patterns?



Case Study - Questions

You know you have some responsibilities for data quality, but what exactly...?

- Are you really responsible for knowing the definitions of all the data fields?
- What if the data contained in the fields are identical? Similar? Not even close? Does it matter?
- Does it matter whether you pulled the data yourself or someone else provided it for your analysis?
- Under what circumstances do you NOT have to worry about data quality?
- If you identify an issue with the data, what can you do?
- What if you never get an answer to your question about the paid loss definitions?
- What – if anything – should you have done differently in reserving?
- **To what extent are you responsible for understanding the data or definitions underlying the external development patterns?**



Case Study - Discussion

Later you finally have time to check on the impact of the difference. Had you used the reserving data in your study, the overall indication would have been the same, but not by coverage or territory. So whatever the difference is, it has an impact on some customers' rates. Now what?

- A. This changes nothing. You still can't make a decision until you actually know what the difference is.
- B. You volunteer to lead a "data documentation" project for your department in conjunction with IT.
- C. Surely this impacts all pricing studies. You ask your boss to delay all further work until you can get to the bottom of this.