In the period since the Office of Inspector General of Health and Human Services issued its updated Provider Self Disclosure Protocol (SDP), organizations have reported that it provides important new guidance in some areas, while being silent in others. The purpose of this white paper is to share some of the practical insights that we have gained from working through SDP implementation challenges with disclosing parties, as they endeavor to follow the specified process for voluntarily identifying, disclosing, and resolving liability for “instances of potential fraud involving the Federal health care programs.”

Although legal guidance is critical in order to determine how to appropriately comport with each disclosing party’s unique circumstances, most providers may need to grapple with some or all of the practical questions enumerated below in their implementation efforts. Not surprisingly, these questions frequently trigger debate and a variety of perspectives from different stakeholders.

1. **How should “Federal” payers be defined?** Under the guidance provided by the previous version of the self-disclosure protocol, disclosing parties typically felt compelled to include Medicare and TRICARE in their definition of “Federal” payers, but were not necessarily required to include Medicaid.

   The updated SDP, however, appears to fairly clearly signal that Medicaid should be included in the study population, even though the reimbursement amounts and coverage determination guidance are often different from Medicare’s.

   Further, our experience is that some disclosing parties have made the principled decision to exclude Medicare and Medicaid HMO plans from the definition of “Federal” payers because of the different economic arrangements associated with these managed care plans. Under this line of analysis, the Federal government is typically not at direct incremental economic risk—and therefore should not be damaged—if the plan reimburses a non-compliant claim on behalf of a covered patient, because that is part of the risk that the managed care plan has assumed when it agreed to accept capitated payments (flat monthly fees) from the Federal government, regardless of an individual patient’s utilization of services.

2. **What patients will be included - inpatient, outpatient facility, or both?** Given the different reimbursement methodologies used for inpatient claims (e.g., MS-DRG, or APR DRG codes billed for a specific diagnosis on a per-discharge basis under the Inpatient Prospective Payment System) versus outpatient claims (i.e., a separate CPT/HCPCS code for each documented, medically-necessary service), some parties may conclude that only a subset of their patients should be included in the study population.

   For example, a provider recently discovered that some of its cancer patients had been administered a specific drug on a prophylactic basis to prevent neutropenia during a time period before coverage for this purpose was specifically allowed. However, since the administration of this drug did not impact the inpatient reimbursements that had been received, the provider concluded that it was appropriate to only include paid outpatient claims in the study population, since those were the payments that could have resulted in economic “damage” to the Federal payers.

---

2 For example, the SDP no longer allows for “spares” in the event that the medical records supporting a sampling unit cannot be located.
4 Formerly known as the Civilian Health and Medical Program of the Uniformed Services (CHAMPUS).
3. **How far back do you go?** Should one automatically go back as far as the applicable statute of limitations allows (typically 6 years), or is there a principled basis to start at a more recent point in time? Depending on what gave rise to the non-compliance—such as an underlying system implementation or process change, or a change in coverage determination guidance—there may be a reason why an alternate date is arguably more relevant to use as the start date of the damages period. Other institutions have advocated for other approaches with varying degrees of success, such as starting with the date when they started deviating from their peers according to PEPPER Report\(^5\) data.

4. **When should you end the period from which you plan to draw your sample?** The updated SDP expects the disclosing party, prior to disclosure, to reasonably confirm that the non-compliant conduct has ended.\(^6\) To reduce the risk of having to make an additional self-disclosure in the same area of liability at some future date, many organizations engage in a robust self-validation process, in order to reasonably confirm that the non-compliant conduct has ended, before finalizing the date range for the study population and selecting a full (i.e., statistical) sample of paid claims. As part of this self-validation process, we are often asked to draw and review a probe sample (e.g., 30 paid claims) to independently assist the disclosing party in demonstrating that the non-compliant conduct has indeed ended.

5. **What is the population of paid claims from which the sample will be selected?** This is typically a combination of which payers and patients are to be included, and the date ranges that the disclosing party determines to be appropriate.

6. **Is there a need to validate the completeness of the study population data before drawing a statistical sample?** It is important to consider reasonably validating the completeness of the study population data extract, as these types of reports are usually “ad hoc” reports by definition, and the disclosing party can gain credibility at this step in the self-disclosure process by being able to reasonably demonstrate that the starting point for the analysis was complete and accurate. There are often many ways to accomplish this with only modest levels of investment. In some circumstances, for example, we have been provided with what was represented to be a complete population of paid claims, and we have reasonably reconciled them to the disclosing entity’s audited financial statements. Often there are other available “control totals” that can be triangulated to reasonably demonstrate that the starting point for the self-disclosure analysis is complete.

7. **What sample size and sampling inputs will be used?** The updated SDP recommends that the estimation of damages consist of a review of either “all the claims affected by the disclosed matter or a statistically valid random sample of claims that can be projected to the population of claims affected by the matter.”\(^7\) The updated SDP specifically states that when using a sample to estimate damages, the sample must contain “at least” 100 sampling units. However, this guidance can be potentially confusing to providers as 100 sampling units may not rise to the level of being a statistically valid sample size, given certain parameters.

Many disclosing parties use a confidence level of at least 90%. First, this was historically the minimum confidence level required by the OIG, and second, using a lower confidence level increases the risk that the disclosing party will end up with an overstated estimate of damages.

So what specifically are the “confidence level,” the “anticipated error rate of occurrence,” and the “desired precision,” in layman’s terms?

---

\(^5\) PEPPERresources.org “is the official site for information, training and support related to the Program for Evaluating Payment Patterns Electronic Report (PEPPER). PEPPER provides provider-specific Medicare data statistics for discharges/services vulnerable to improper payments. PEPPER can support a hospital or facility’s compliance efforts by identifying where it is an outlier for these risk areas. This data can help identify both potential overpayments as well as potential underpayments” (accessed April 14, 2014).

\(^6\) Updated OIG’s Provider Self-Disclosure Protocol, issued on April 17, 2013.

\(^7\) Updated OIG’s Provider Self-Disclosure Protocol, issued on April 17, 2013.
**Confidence level:** This is a numerical measure of how confident one can be that the results of the sample are consistent with the results that would have been obtained if one had tested the entire population as a whole. A higher confidence level can typically cause a larger random sample to be generated than if a lower confidence level were selected.

**Anticipated rate of occurrence (used when performing attribute sampling):** The average number of errors based on known historical data or informed judgment. In the absence of such data, a 50% anticipated error rate is the most conservative assumption to use.

**Desired precision range:** Also known as the “width of the confidence interval,” this is a measure of how “precise” or “tightly focused” the observed error rate is. For example, if it is determined that the observed error rate in the random sample of paid claims is 30%, and a 90% confidence level and 10% precision range were used in designing the sample, then the disclosing party can conclude that it is 90% confident that the true error rate in the population as a whole is between 27% and 33% (i.e., 30% observed error rate, plus or minus 10%). Historically, the OIG required a minimum precision of plus or minus 25% with 90% confidence. The updated SDP no longer states a minimum required precision range.

8. **Does the disclosing party have to separately pay back payments received for unsupported claims that were identified during the probe sample?** If these claims were included in the study population from which a statistically valid sample was later drawn, and if a repayment was made based on that statistical sample, then it would not be appropriate to separately repay improper claims from the probe sample as that would result in a double payment to the government.

9. **When the random statistically valid sample is drawn, should the disclosing party analyze it to demonstrate that the sampled units are reasonably representative of the population as a whole?** If the error rate results are different from what key stakeholders had originally anticipated, a typical reaction is to ask if the sample was representative of the study population as a whole. By performing this analysis as part of the SDP process, it can be very helpful in helping to confirm credibility around the outcome of the review process.

10. **Can the disclosing party present an offset (a “but-for” billing) in its damages analysis?** When calculating and disclosing damages estimates to the government, our experience is that disclosing parties often consider presenting a proposed “offset” measure to gross damages by quantifying what level of reimbursement would have been supported in the “but-for” world, assuming that the relevant facts and medical records support this. While the government may not accept the proposed “but-for” reimbursements as an offset to estimated total damages, many disclosing parties have found that the government has been open to this approach as a matter of equity, as long as it is disclosed and presented in a transparent manner, and if consistent with the underlying facts and circumstances of the situation. Further, because the OIG has stated that its “general practice in [civil monetary penalty] settlements of SDP matters is to require a minimum multiplier of 1.5 times the single damages,”8 every dollar that single damages can appropriately be reduced will typically save the disclosing party at least $1.50.

    It is important to note that the updated SDP prohibits a “…reduction, or ‘netting’ for any underpayments discovered in the review.”9 However, if a significant number of underpayments are noted during the review process, some disclosing parties will opt to include this information in the SDP report, as it could be a relevant consideration in their negotiations with the government.

---

8 Updated OIG’s Provider Self-Disclosure Protocol, p. 2.
11. **Can alternate sample units be selected if supporting information can’t be found?** Even though the RAT-STATS program continues to technically allow for the selection of alternates, the updated SDP does not provide for the use of alternates. It states instead that missing sample items should be treated as errors, citing Federal health care program rules which require the retention of supporting information for submitted claims.

12. **What information is the disclosing party relying upon in addition to the medical records?** In some cases, documentation that could potentially help support a claim may reside outside the patient’s chart. For example, we noted a disclosing party recently decided to take the position—for self-disclosure purposes only—that physician orders that happened to have been maintained in the lab due to how initial workflows were set up could appropriately be considered when assessing whether the underlying claim was compliant.

13. **Which coverage standards to apply?** When approaching the claims review, applying Medicare rules to all Federal payers can expedite the review process, but it can also result in overstating the disclosing party’s error rate. However, applicable historical Medicaid coverage rules are often not as readily assessable as Medicare’s. Therefore, an important time-saving step in the review process can be to first select the random sample of paid claims, and then identify which rules, from which years and geographies, need to be located in order for the reviewers to be able to use the most relevant Medicaid rules for the applicable randomly-selected paid claims.

14. **Do you select a separate random sample for each Federal payer?** Historically, the disclosing party would often define a separate sampling frame and draw a separate statistical sample for each Federal payer. However, the updated SDP appears to allow for a disclosing party to draw one combined sample from a combined population of all Federal paid claims of interest, and then allocate a pro rata share of the resulting consolidated damages to each payer in a systematic and rational way: “if the disclosing party’s Federal payor mix is 60 percent, the disclosure should break down how the Federal health care programs make up that 60 percent, such as 40 percent Medicare, 10 percent Medicaid State A, 5 percent Medicaid State B, and 5 percent TRICARE.”

15. **Include or exclude private payers when investigating paid claims for compliance?** Voluntary self-disclosures can often become public (e.g., through a settlement announcement, or a Corporate Integrity Agreement). Accordingly, and for a variety of reasons—such as reaffirming the tone at the top, mitigating escheat law compliance risks, and/or because some States allow whistleblowers to lodge a complaint in situations where private payers have been overcharged—some disclosing parties decide to include fee-for-service private payers in the claims review process. This allows the institution to get comfort around its overall exposure on a particular type of at-risk paid claims, and to be able to take the lead on reaching out to private payers and getting closure on the exposure at the same time that it is resolving exposure with its Federal payers.

16. **What should the sampling unit be (e.g., a claim, a claim line, a patient)?** The answer can be impacted by a variety of factors, such as the type of service of interest, or how the data are available. For example, charge data may be posted at one level of granularity, while subsequent payment data may be applied at a less granular level, and so it is important to discuss this carefully up front.

---

10 RAT-STATS is the statistical sampling software promulgated by HHS OIG, and the primary statistical tool for HHS OIG’s Office of Audit Services. The use of this software is “strongly” recommended when selecting the random numbers for the sample (see Updated OIG’s Provider Self-Disclosure Protocol, p. 8). It is currently available free of charge on the OIG website.

11 Updated OIG’s Self-Disclosure Protocol, p. 10.
Contacts

For a deeper discussion, please contact:

Kristafer Swanson
PwC
(312) 298-6195
kris.swanson@us.pwc.com

Patricia Pelaez Kawashima
PwC
(312) 298-5564
patricia.a.pelaez.kawashima@us.pwc.com