José Ramón Martínez
CoGerente de Asesoría y Validación, S.L. (ASYVAL)

DATA INTEGRITY: EN QUÉ CONSISTE

En este artículo se pretende aclarar los aspectos principales de la integridad de datos como una interpretación de las normativas vigentes, fundamentada en las diversas guías editadas y en mi propia experiencia.

En los últimos años las observaciones y deficiencias reportadas respecto a la integridad de datos ha ido en aumento, hasta el punto que las entidades regulatorias se han alarmado por el creciente incumplimiento de este punto contemplado en normas tales como la 21 CFR PArt 11 o el Anexo 11 de las GMPs Europeas. Dado que se trata de un problema de calidad importante, las entidades regulatorias tenían que tomar medidas, y por ello en 2015 la MHRA publicó una primera guía para ayudar a focalizar el problema y garantizar la información de los datos. A partir de esta primera guía, se han ido sucediendo otras, tales como:

- *FDA Draft Guidance for Industry: Data Integrity and Compliance with cGMP, April 2016, US Food and Drug Administration (FDA)*
- *MHRA GxP Data Integrity Definitions and Guidance for Industry, Draft version for consultation, July 2016*
- *WHO Technical Report Series, No. 996, Annex 5: Guidance on Good Data and Record Management Practices, World Health Organization (WHO), 2016*
- *PIC/S Draft Guidance: PI 041-1 (Draft 2) Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments, August 2016, Pharmaceutical Inspection Co-operation Scheme (PIC/S)*
- *ISPE GAMP® Guide Records and Data Integrity, April 2017*

Todas estas guías persiguen el principio de implementar estrategias útiles y eficaces, que permitan reducir los riesgos relacionados con la integridad de los datos.

Por todo ello, la integridad de datos se ha convertido en una prioridad para los inspectores europeos y americanos que velan por la obligatoriedad de la industria farmacéutica para asegurar que los datos que generan sean fiables.

Objetivos

En este artículo se pretende aclarar los aspectos principales de la integridad de datos como una interpretación de las normativas vigentes, fundamentada en las diversas guías editadas y en mi propia experiencia. El alcance del artículo no sustituye a la lectura de las guías editadas, sino que constituye una lectura previa que pretende ayudar a su comprensión.

Principios

Los principios a tener en cuenta para entender las guías son:

- Están dirigidas a todas las empresas donde se requiere una integridad en los datos, como farmacéuticas, biológicas, producto sanitario, cosméticas, alimentarias, etc…
- Están pensadas únicamente para datos regulados, que son aquellos que contienen información utilizada para un propósito regulado o para apoyar un proceso regulado.
- Sostienen que los datos sujetos a las normas de calidad sólo son los datos primarios, es decir, aquellos que se utilizan para actividades reguladas. Estos datos, en principio, son los obtenidos de forma original y son únicos.
- Tienen en cuenta metadatos que describen atributos y proporcionen contexto y significado a los datos. Normalmente describen la estructura, los elementos, las interrelaciones y otras características de los datos.
- Contemplan registros regulados, como un conjunto de datos y sus metadatos regulados con un propósito, contenido y significado específicos y requeridos por normas GxP.
- Consideran que el soporte para datos y registros regulados puede ser en papel, de forma electrónica o situaciones híbridas (aunque se fomenta un alejamiento de las situaciones híbridas, siempre que sea posible).

Los puntos principales para asegurar un correcto cumplimiento de las normas que contienen la integridad de datos, y por tanto seguimiento de las guías editadas, son:

- Establecer una política corporativa de integridad de datos dentro del Sistema de Gestión de Calidad (SGC)
- Disponer de un sistema de gestión de riesgos para la integridad de datos
- Establecer un sistema de razonamiento crítico para los datos regulados
- Disponer de un sistema de gestión de datos que contemple todo el ciclo de vida de los datos regulados
- Verificar el cumplimiento de requerimientos de integridad de datos en los datos regulados

Fig. 1 - Esquema ALCOA

Requerimientos de integridad de datos

Los requerimientos que deben cumplirse para asegurar la integridad de datos están contenidos en el acrónimo ALCOA (*Attributable*, Legible, *Contemporaneous*, Original y *Accurate*):

Atribuible: El dato debe poderse asignar a la persona o sistema que lo genera. Asimismo, identificar si la generación del dato ha sido por creación o modificación.

Legible: El dato debe ser comprensible y permanente. Asimismo debe poderse acceder a él durante todo su ciclo de vida. Este dato ha de ser original y sus modificaciones no deben esconderlo.

Contemporáneo: El dato debe registrarse u observarse en el momento en que sucede la acción.

Original: Debe conservarse el primer registro del dato con su contenido y significado.

Accurate (preciso): El dato no debe contener errores y las correcciones deben quedar documentadas.

Suelen añadirse también los siguientes criterios para formar el acrónimo ALCOA++++

+ Completo: Debe disponerse de todos los datos y metadatos relevantes.
+ Consistente: Los datos deben documentarse según buenas prácticas de documentación y deben disponer de marcas de fecha y hora en la secuencia esperada.
+ Duradero: El dato debe estar registrado de forma permanente durante el periodo de retención.
+ Disponible: El dato debe estar a disposición ante cualquier revisión, auditoría o inspección durante todo el periodo de retención.

Política de integridad de datos

Las empresas sometidas a las normas que incluyen la integridad de datos deben implementar estrategias significativas y efectivas para gestionar los riesgos de integridad de datos basados en la comprensión de sus procesos y la gestión del conocimiento de tecnologías y modelos de negocio.

La integridad de los datos y la gestión de datos se deben incluir en el Sistema de Gestión de Calidad (SGC). Las actividades definidas deben estar dirigidas por la alta dirección. En este punto debe incluirse:

- Un plan para la mejora continua en los programas corporativos para la integridad de datos
- Un inventario de los sistemas que generan y retienen los datos, así como la capacidad de dichos sistemas. También se debe disponer de un inventario de documentos dentro del SGC.

- Un sistema de autoevaluaciones siguiendo procedimientos definidos para determinar el grado de control de la integridad de datos.
- Medidas de control comportamental, procedimental y tecnológicas para asegurar la integridad de datos.

Si se realiza una revisión del SGC frente a los requerimientos de integridad de datos se pueden identificar los controles que se deben realizar. El SGC es correcto si pueden darse respuesta a las siguientes preguntas:

- ¿Existe un proceso adecuado en el SGC para la prevención, detección, etc. de un fallo en la integridad de datos?
- ¿Los requerimientos de integridad de datos son claros en el SGC?
- ¿La generación y revisión de datos están claramente definidas en el proceso?
- ¿Existen los controles adecuados a lo largo del ciclo de vida de los datos?

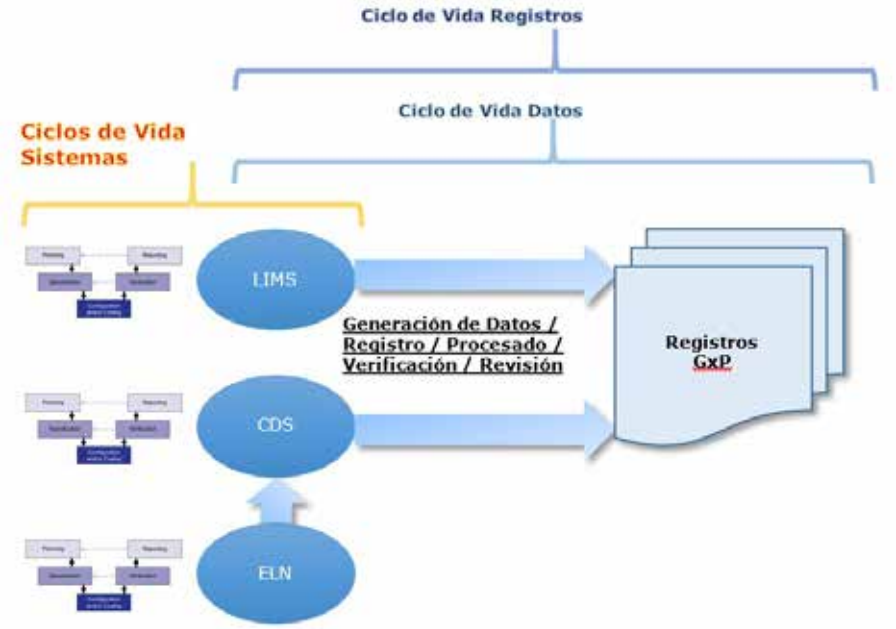

Fig. 3 - Relación ciclos de vida.

Ciclo de vida de los datos regulados

Las fases por las que atraviesan los datos primarios son:

- Creación y captura
- Registro inicial de datos mediante su procesado (incluida la transformación o la migración)
- Revisión, presentación de informes y uso
- Retención y recuperación
- Destrucción

La integridad de los datos debe garantizarse durante todo el ciclo de vida de los datos.

Este ciclo de vida debe distinguirse de los ciclos de vida de sistemas informatizados y del ciclo de vida de los registros. El ciclo de vida de los sistemas informatizados es totalmente independiente de los otros dos (como se muestra en el ejemplo de la figura 3) y el ciclo de vida de los registros es la suma de los ciclos de vida de los datos que lo constituyen:

Creación y captura de datos

Los datos pueden crearse por entrada manual o desde un dispositivo o sistema. La integridad de datos puede verse comprometida por fallos en la fiabilidad, la exactitud, la precisión, el contenido y el significado. Las medidas para reducir estos riesgos deben responder a un análisis de riesgos, pero en general debe considerarse:

- Todos los datos (incluso si están excluidos) se deben recoger y conservar para el uso previsto, con la exactitud, integridad, contenido y significado adecuados. Asimismo, deben estar disponibles para su revisión.
- Los riesgos para la integridad de los datos están influenciados por el grado en que potencialmente pueden ser manipulados (accidental o intencionadamente).

Procesado y registro de datos

Los datos pueden requerir de procesados definidos (cálculos y algoritmos) para obtener y presentar información en el formato requerido. La integridad de datos puede verse comprometida, básicamente, por la fiabilidad de éstos. Como en la etapa anterior, las medidas para reducir estos riesgos deben responder a un análisis de riesgos y, en general, considerarse:

- Los datos que se procesan y/o reprocesan deben seguir procesos definidos y verificados, para obtener y presentar información en el formato requerido
- El impacto del proceso de datos sobre la calidad del producto y la seguridad del paciente variará según el producto y el proceso

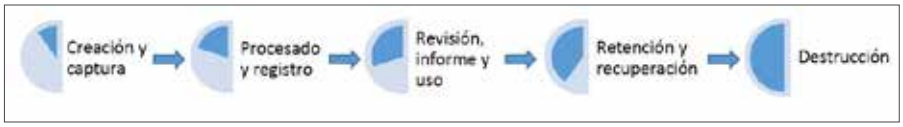

Fig. 2 - Ciclo de vida de datos.

Revisión de datos, informe y uso

Durante esta fase los datos se utilizan para tomar decisiones. Las actividades contempladas incluyen la revisión, la presentación de informes y el uso de datos, de acuerdo con procesos definidos, verificados y con procedimientos aprobados:

- Revisión de datos
  La revisión de datos debe determinar si se ha cumplido con especificaciones predefinidas, objetivos, límites o criterios. Deben considerarse todos los datos, teniendo en cuenta que los datos pueden ser almacenados en diferentes lugares. Esto incluye datos atípicos, sospechosos, rechazados, no válidos o eliminados, junto con cualquier justificación. Las empresas deben documentar los procesos de revisión realizados y el procedimiento seguido para la revisión.
- Revisión del Audit Trail
  Las empresas reguladas deben establecer un proceso documentado para la revisión de los registros de Audit Trail, como parte de la rutina de revisión/aprobación de datos. Por lo general, la lleva a cabo un responsable del área operacional que ha generado los datos y una persona que entienda el proceso y el impacto de las acciones registradas. Son un medio eficaz para verificar que los cambios están realizados por usuarios autorizados y para detectar posibles problemas de integridad de datos.
- Informe de datos
  Los procedimientos de notificación de datos deben garantizar la coherencia y

la integridad de los resultados. Para reducir los riesgos sobre la integridad en los informes de datos pueden tomarse medidas como:

- Prestar una atención especial cuando los usuarios puedan influir en el informe de datos.
- Evitar métodos e informes mal diseñados o definidos, ya que pueden ocasionar posibles decisiones erróneas.
- Los resúmenes de datos deben incluir una verificación documentada contra los datos originales.

- Distribución de datos

  Deben tenerse en cuenta los riesgos asociados a:
  - Distribución de los datos a personal autorizado.
  - Acceso a los datos por parte del personal autorizado.
  - Envío de datos a otros sistemas del proceso.

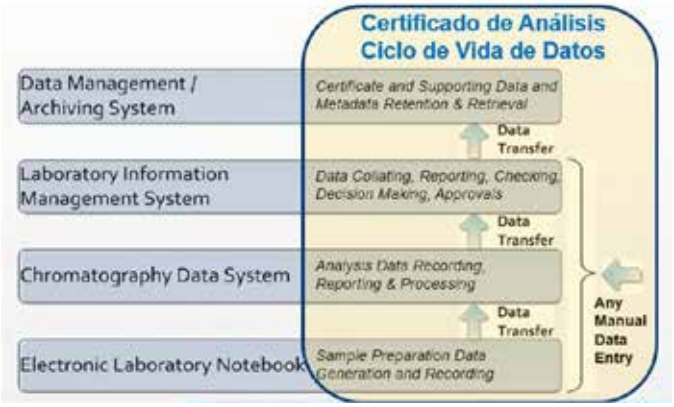

Fig. 4 - Ejemplo de Ciclo de Vida de datos para un certificado de análisis.

Retención y recuperación de datos

La retención de datos es un término amplio que hace referencia a la forma en la que se conservan los datos (electrónicos o papel) y sus copias de seguridad. Los datos que deben retenerse son aquellos necesarios para mantener el contenido y el significado GxP.

Para reducir los riesgos sobre la integridad en la retención y recuperación de datos pueden tomarse medidas como:

- El archivo con el propósito de revisión o investigación debe realizarse utilizando un proceso definido, verificado y aprobado.
- Las copias de seguridad deben realizarse de forma periódica de todos los datos pertinentes, incluidos los metadatos, según un proceso documentado. Los controles serán proporcionales a los de los datos originales.
- Debe realizarse una verificación para asegurar que los datos almacenados puedan recuperarse después de cambios en el sistema (hardware o software), preservando el contenido y el significado requerido.
- Cuando se utiliza un tercero para la retención y recuperación deben tenerse en cuenta los riesgos asociados y establecer un acuerdo formal para definir las funciones y responsabilidades respectivas.
- Cuando se retira un sistema o éste no puede conservarse, debe tenerse en consideración la forma en la que los datos podrán seguir almacenados y recuperarse durante el periodo de retención.

Destrucción de datos

Cuando los datos superan el periodo de retención se eliminan de acuerdo con un proceso definido y procedimientos aprobados. Debe asegurarse que los datos originales se eliminan, así como las copias distribuidas.

Sistema de gestión de riesgos para la integridad de datos

El sistema de gestión de riesgos utiliza las herramientas habituales, según el enfoque ICH Q9 y GAMP 5, pero aplicado a una gestión holística de los datos, es decir mediante una comprensión completa del flujo de datos de la empresa, de forma que se incorporen medidas de control de riesgo.

Deben considerarse los riesgos asociados a los sistemas (automatizados y manuales) y a los usuarios.

Las medidas de control de riesgo para los sistemas automatizados se distribuyen en:

- Controles de registros y datos: Gestión de accesos y seguridad; Copias de seguridad y su restauración; Continuidad del negocio; Audit Trail; Software y Hardware; Políticas y procedimientos; Formación y experiencia; etc.
- Controles de firma electrónica: Métodos para asegurar que el ID es único; Prevención de la eliminación de información relacionada con la firma después de haberse firmado; Métodos biométricos; Firma digital; Técnicas para garantizar la vinculación entre el registro y la firma; etc.

Los sistemas híbridos aportan riesgos adicionales en la interface entre el manual y el sistema automatizado.

Entre las medidas de control de riesgo para los factores humanos deben considerarse las diferencias culturales, el error humano, problemas de comprensión, de conocimientos, de motivación y comportamiento.

Sistema de razonamiento crítico

El razonamiento crítico es un proceso sistemático, racional y disciplinado de evaluar la información desde una variedad de perspectivas para producir una respuesta equilibrada y bien razonada. El razonamiento crítico permite la interpretación efectiva de datos y situaciones, evitando prejuicios personales, suposiciones y otros factores. Ayuda a conseguir la integridad de datos, pues permite:

- Identificar brechas en los procesos de datos.
- Desafiar la efectividad de los controles técnicos, procedimentales y de comportamiento.
- Identificar y evaluar riesgos.

Sistema de gestión de datos

El sistema de gestión de datos definido en las guías es la suma total de las medidas adoptadas para asegurar que los datos son registrados, procesados, retenidos y utilizados para asegurar un registro completo, consistente y preciso durante todo el ciclo de vida de los datos.

La gestión de datos abarca a las personas, los procesos y la tecnología necesarios para lograr una gestión consistente, precisa y efectiva. Proporciona la estructura dentro de la cual se toman las decisiones apropiadas sobre asuntos relacionados con los datos, de acuerdo a modelos, principios, procesos y autoridad definida. También puede considerarse como un enfoque de control y garantía de calidad para aplicar rigor y disciplina al proceso de administración, uso, protección y mejora de la información organizacional.

Las actividades de gestión de datos deben estar integradas en el SGC. Puede ser que sea necesario un programa corporativo específico para la integridad de datos.

Se puede fijar un rol de nivel ejecutivo como *Chief Data Officer* (CDO) para la gestión de los datos.

La gestión de datos debe ser para todos los datos de la empresa, por lo que también debe ser holística, proporcional e integrada.

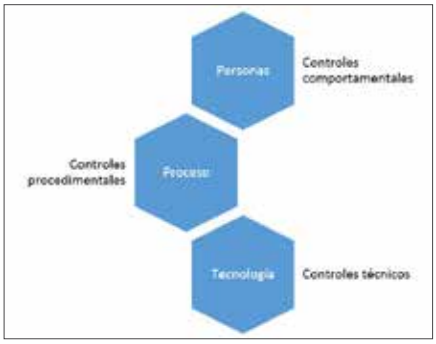

Fig. 5 - Alcance de la Gestión de Datos y sus controles asociados

Fig. 6 - Marco de la Gestión de Datos.

Elementos del marco de gestión de datos

Los elementos que forman parte del sistema de gestión de datos son:

- La organización de gestión de datos.
- Los controles sobre los datos.
- Los procesos que aseguran la integridad de datos.
- Las infraestructuras que soportan la gestión de datos.

Factores humanos en la integridad de datos

Tanto las diferencias de culturas corporativas (forma en que opera cada empresa) como las diferencias de cultura local (geográfica) pueden influir en la gestión del riesgo, facilitando o dificultando su gestión.

Controles

Hay 3 tipos de control para conseguir un nivel aceptable de integridad de datos:

- Comportamental
  La alta dirección debe asegurarse de que el personal no está sujeto a conflictos/presiones de tipo comercial, político, económico… que puedan afectar a su trabajo y a la integridad de los datos.
- Procedimental
  Las empresas reguladas deben implementar procedimientos para disminuir el riesgo potencial asociado a la integridad de datos.
- Técnico
  Los controles técnicos son los que se aplican a los sistemas informatizados de gestión de datos (cualquier actividad dentro del ciclo de vida de los datos), para reducir el riesgo contra la integridad de los mismos. Es imprescindible que los sistemas informatizados GxP relevantes se validen para su uso previsto. En esta validación se comprobará la existencia y efectividad de los controles técnicos diseñados.

Modelo de madurez de la integridad de datos

Las empresas reguladas deben incluir en el diseño de la tecnología, antes de su adquisición, la integridad de los registros y los datos. Los sistemas adquiridos deben poder configurarse para proporcionar la integridad adecuada de los datos.

En la Guía GAMP "*Records and Data Integrity*" se introduce el modelo de madurez de integridad de datos (descrito en el Apéndice M2). Es una representación simple del estado de los elementos esenciales de la empresa respecto a la integridad de datos.

Para categorizar el nivel de madurez se proporcionan unas tablas de ayuda que contemplan: cultura, gestión y organización, plan estratégico y programa de integridad de datos, regulación, ciclo de vida y ciclo de vida de los procesos de apoyo. Las tablas tienen la siguiente estructura:

Puede ser que una misma empresa se encuentre en distintos niveles, debido a que cada área/proceso se encuentra en un nivel distinto.

Conclusión

Aunque es cierto que las normas que incluyen el concepto de integridad de datos hace tiempo que existen, el enfoque desarrollado por las nuevas guías como respuesta a los últimos escándalos respecto a la fiabilidad de los datos, introduce nuevas rutinas organizativas y actividades en las empresas bajo regulaciones de calidad.

Las empresas reguladas deben:

- Considerar la integridad de los datos y registros según requerimientos ALCOA.
- Establecer y mantener políticas y procedimientos robustos para todos los aspectos del ciclo de vida de los sistemas de gestión de datos.
- Establecer un sistema de gestión de riesgos holístico que desafíe la integridad de datos.
- Utilizar un proceso de razonamiento crítico para evaluar información y permitir la mejora continua.
- Establecer un sistema de gestión de datos con los controles adecuados (comportamentales, procedimentales y técnicos) ◂◂

Maturity Area	Maturity Factors	Maturity Level Characterization				
		Level 1	Level 2	Level 3	Level 4	Level 5
Culture						
Data Integrity Understanding and Awareness	Awareness of the importance of data integrity and understanding of data integrity principles	Low awareness, limited to SMEs and specialists	General awareness of the topic, but not fully reflected in working practices	Principles reflected in working practices, but not consistently applied	Data integrity principles fully incorporated and applied in established processes and practices	Formal ongoing awareness program, proactively keeping abreast of industry developments

Fig. 8 - Fuente: Guía GAMP *Records and Data Integrity* de ISPE.

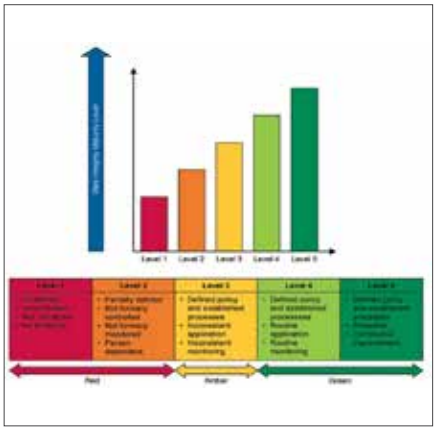

Fig. 7 - Fuente: Guía GAMP *Records and Data Integrity* de ISPE.

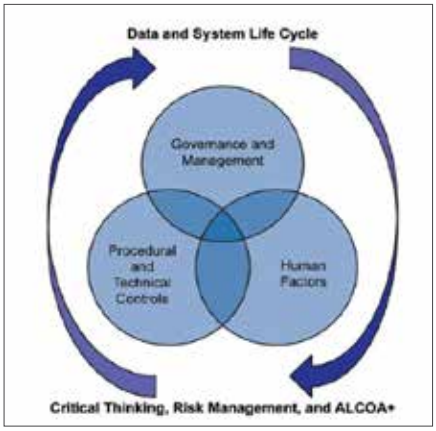

Fig. 9 - Fuente: Guía GAMP *Records and Data Integrity* de ISPE