# 

*Version 3.0*

Data Management Guidelines

DMG 13 – Master Data Management

Table of Contents

[0](#_Toc79358082)

[1. Why do we need Master Data Management? 2](#_Toc79358083)

[2. What is Master Data Management? 7](#_Toc79358084)

[3. MDM Implementation Styles/Patterns 8](#_Toc79358085)

[3.1 Registry 8](#_Toc79358086)

[3.2 Consolidating 9](#_Toc79358087)

[3.3 Co-Existence 9](#_Toc79358088)

[3.4 Transaction/Centralized 10](#_Toc79358089)

[4. The Standards relevant to MDM 11](#_Toc79358090)

[5. MDM Best Practices 12](#_Toc79358091)

[6. MDM Roles and Responsibilities 13](#_Toc79358092)

[7. MDM Architecture 14](#_Toc79358093)

[8. Useful References 16](#_Toc79358094)

[9. Document Administration 16](#_Toc79358095)

# Why do we need Master Data Management?

Incomplete, incorrect, outdated, redundant, inconsistent, or misaligned Master Data (see section 2) is likely to have adverse impacts on safety, business process efficiency, business intelligence, and regulatory compliance. Conversely, accurate and up to date Master Data is critical to the economic, customer service and safety performance of National Grid. Master Data is also a key enabler of standardisation and efficiency in business operations and in the tactical and strategic use of the management information.

In National Grid, our technology landscapes have evolved over time and we generally have multiple copies of data stored in varying formats in different systems or spreadsheets with poor control of those copies and often without a general agreement on which is the master authoritative source to be used. The maintenance of this duplicate data is performed separately and in isolation and often in differing parts of the organisation. To prevent discrepancies between the data copies on different applications, it is vital to ensure appropriate mechanisms are introduced to ensure consistency of the Master Data is maintained.

The same or similar Master Data is often used and maintained by different departments and people with different perspectives therefore we need a consistent Master Data design across all those areas. Traditionally we have developed solutions without the use of standard reference tables (lookup tables) which undermines our ability to manage, integrate and exploit our data.

This poor Master Data quality can and does have significant adverse effects on the overall performance of National Grid:

* Bad Master Data causes expensive disruptions in operations and processes, such as:
  + Incorrect planning due to the wrong address leads to re-scheduling, additional follow up work and reduced customer satisfaction.
  + Inconsistent Master Data in work management causes incorrect financial coding of work requiring manual correction.
* Bad Master Data can lead to increased operational risk, such as:
  + Problems with traceability of components and component types
  + Incorrect assessment of asset management requirements based on inconsistent asset and work data across multiple applications
* Bad Master Data undermines business information, such as:
  + Inconsistent business information is a key issue during regulatory reporting processes
  + Manual off-line realigning of key information results in a high cost of data preparation and analysis, reduced timeliness and decreased accuracy.
  + Discussions and doubts about the accuracy of data cause many business discussions to revolve around data arguments and disagreements rather than the business issues.

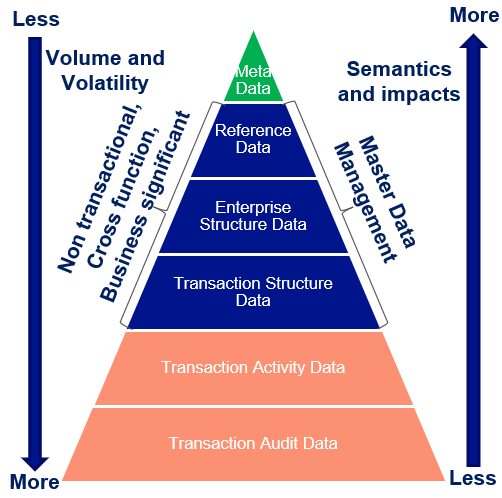
The key reasons for proactively addressing the management of multiple copies of data are:

* There are overheads associated with maintaining multiple copies of data in multiple repositories (whether in centrally controlled systems or in spreadsheet copies). If we can reduce the level of duplication of data sources, we can reduce the maintenance overheads and achieve a more efficient and accurate level of data control.
* Having the option of sourcing reports from alternative data sources creates unnecessary problems. Significant time is consumed when resolving reporting discrepancies and inconsistencies when it is unclear which data sources have been used, and which are the right master sources to use.
* We need to ensure people understand and use the ‘right’ data sources for operational purposes and reporting – thus reducing the opportunity to create reporting discrepancies.
* A critical factor in managing the use of copies of data is to ensure users access slave copies of data at the right time (i.e. recently synchronised with master) – running monthly reports the day before the monthly refresh should be challenged!

Implementing Master Data Management (MDM) will enable National Grid to realise significant benefits by:

* **Improving operations.** Accurate asset location information will significantly improve efficiencies and reduce costs associated with some of the key business processes.
* **Increasing customer satisfaction.** Effective Master Data management can provide a unified view of the customer by leveraging customer data across multiple platforms. In addition, Master Data management is a key prerequisite to reaping benefits from innovative use of customer interaction platforms, from a monthly bill to the self-service internet portals.
* **Increasing insight and transparency.** Clean and accurate data improves organisation’s ability to measure and interpret organisational performance and serves as a key enabler for improved business decisions.
* **Ensuring compliance.** Well implemented Master Data management programmes can help utility companies address the critical business concern of regulatory compliance. Master Data management will address a broad range of compliance and governance issues, including Sarbanes-Oxley, IFRS and Global Trade Compliance. Benefits of improved Master Data quality and the implementation of the correct processes in these areas can yield significant savings in audit related costs.

The other type of data that this guideline covers is Reference Data. This data type is frequently over looked when business entities are implementing Master Data governance and change control processes as it is much lower in volume and frequency of change than traditional Master Data. However, it can have significantly more far-reaching consequences (as shown by the diagram below) and therefore is at least as important to manage appropriately.



For example: One mistake on a transaction will only effect that transaction. One mistake on a bit of Master Data will affect every transaction that uses the Master Data item. One mistake on a bit of reference data will affect every bit of Master Data that uses it and in turn every bit of transactional data. Thus, the error is compounded.

In summary, successful Master Data Management relies on a high-quality design of the logical data structures, of the use of standard data objects, of the governance and data maintenance processes and of the design of processes and interfaces which keep consistency across multiple applications.

The scope of Master Data management here is to ensure master source data and related copies are successfully synchronised and aligned on a frequency appropriate to their business use. Where we have replicated data and no master source is formalised, the requirement is to establish the best source to deem master and then establish master-slave relationships between copies of data.

It is therefore advisable, during any data change project, to dedicate significant attention to:

* The target system Master Data structures
* Master Data maintenance processes

In order to understand the importance of Master Data, then it needs to be seen in context of other data types. All data can be assigned to one of the following definitions:

|  |  |
| --- | --- |
| Data Classification | Description and Example |
| Metadata | Metadata is data that is used to describe data. i.e. data which describes the structure and taxonomy of information. Defining metadata correctly is critical to the ability to store and manage the data at a practical level. Correctly done it can help support data quality by predefining data capture parameters.  Example:   * A CUST\_ID field is to store Customer ID and has a size of 10 alpha numeric characters and it cannot be blank (NOT NULL). |
| Master Data | **Reference Data -** Reference data is data that defines the set of permissible values to be used by other data fields. Reference data gains in value when it is widely re-used and widely referenced. Typically, it does not change overly much in terms of definition (apart from occasional revisions).  Example:   * Type of asset, Status codes, Reason codes, Type codes, Classification codes, Country, State, Zip/Postal codes   **Enterprise Structured Data** – Enterprise structured data defines the hierarchies within the enterprise.  Example:   * National Grid Organization Hierarchy like Company, Business Unit, Division, Department etc   **Transaction Structure Data** – These are the ‘nouns’ of the business – they are the things that will participate in events. These will be referenced to perform transactions. This is data about the business entities that provide context for business transactions  Example:   * Asset, Customer, Customer Profile, Worker, Location, Position etc. |
| Transaction Data | **Transaction Activity Data -** Transaction activity data is data describing an event (the change as a result of a transaction) and is usually described with verbs.  Transaction data always has a time dimension, a numerical value and refers to one or more master data  Understanding data as being ‘transactional data’ will require us to understand the volumes involved and to determine the appropriate storage requirement.  Example:   * Customer Billing and Payment, Work Request, Work Order, Worker Learning and Salary/Payment.   **Transaction Audit Data -** Transaction audit data is data that defines logs of transactions executed to bring about a process flow  Example:   * Create, update, delete and read audit log records |

The key area covered by this document concern our management of:

* **Master Data:** we need to actively manage our current legacy of replicated data and improve controls such that the level of replication is reduced as opportunities arise. If we can encourage more REUSE of Master Data, the overheads in managing replication can be reduced. Also, we need to establish common reference tables which are shared across multiple systems. If we can establish shared and common reference tables, their REUSE will have real business benefit as described in the section below.

There are less concerns currently with our overall management of ‘transaction data’– if there are issues with these, they will be addressed on a localised basis.

While defining Master Data can be fairly straightforward, in practice there can be some confusion on whether an object should be considered and managed as Master Data or whether it should be treated as another (e.g. transactional) data type. Some general indicators that may help to decide whether the object should be considered and managed as Master Data are listed below:

* **Behaviour** – Master Data interacts with transactional data within the organization. Most commonly when one describes a business interaction Master Data is captured within a noun while transactional data type is captured by a verb in the sentence (e.g. “*Vendor* sells a *part*, and a *partner* delivers *materials* to a *location*”)
* **Cardinality** – in general, as the number of records in data set decreases the need for data elements to be considered and managed as Master Data decreases (e.g. if a company has five customers it does not need to deploy a MDM solution and full governance structure to manage this data). Though it is still Master Data and should have the appropriate processes and controls.
* **Volatility** - Master Data tends to be less volatile than transactional data.
* **Re-use** – Master Data objects are typically re-used in the course of businesses operations. i.e. shared or duplicated across systems.
* **Value and complexity** - The more valuable the data object is to the business and the more complex it is (in terms of attribute number, cross attribute validations etc.) the greater benefit there is when it is to be treated as a Master Data element within an MDM solution.

# What is Master Data Management?

In National Grid we define Master Data Management (MDM) as:

“Master Data Management is the on-going reconciliation and maintenance of Master Data”.

Master data management is a set of policies, standards, processes, governance, stewardships, and tools that combined create a commonly trusted, consistent, accurate, and controlled set of “master data” for critical business from across internal and external data sources and applications.

Master data management is used for building reliable, accurate, non-duplicative master data that National Grid business can use for effective decision making and digital transformation. Core master data entities include Customer, Workforce, Asset, Location, Product, Vendor, Reference etc.

# MDM Implementation Styles/Patterns

MDM implementation styles are crucial for successful MDM solution deployment. They play a key role in architecting the MDM system. Right MDM implementation style improves the quality of your master data and enhances the consistency and managed use of this information in what is often a complex and somewhat tangled environment. It also helps you support the operational environment/decision-making environment, push clean data back into existing systems, build API fabric, enable demographic distribution, and other unique MDM requirements. Different implementation styles have come to the fore to address different business needs. There are four most common MDM implementation styles/patterns. Depends on the complexity, National Grid may have different implementation patterns for different master data domain.

## Registry

In this style, build an MDM index using only limited key identifier master data fields. Most of the data stays in source systems. MDM keeps track of where each master data element is stored in the source system. When a request is made for the master data set, MDM returns the full consolidated set after getting appropriate master data from the appropriate source systems.



**Key highlights:**

* De-duplication of records
* Employs data cleansing – match/merge
* Data is not sent back to source systems
* Governance required to ensure reliable golden record

**Pros:**

* Less intrusive as no impact to source systems
* Quick and inexpensive to set up

**Cons:**

* Complex to build 360 view if more source systems
* Require most MDM configuration
* Requires retrieving master data from the source systems

**Possible Use Cases for National Grid:**

* When there is only very limited (one or two) source systems for system of record/truth for the master data

## Consolidating

In this style, build an MDM centralized repository with curated consolidated master data, allow data stewards to correct data in the MDM tool, and cleansed master data are not sync up with source systems.



**Key highlights:**

* Data consolidated into the MDM repository from multiple sources
* Data is cleansed to create single source of truth, high data quality system
* Data Stewards can clean data in MDM
* MDM Data is not sync up with source systems

**Pros:**

* Less intrusive as no impact to source systems
* Quick and inexpensive to set up
* MDM is a repository for Golden master record
* Standard MDM Configuration

**Cons:**

* Possible to have data discrepancies between source systems and the MDM tool
* Requires manual updates in source systems for data quality issues

**Possible Use Cases:**

* When there are multiple source systems and source systems are not able to take data updates automatically from other systems

## Co-Existence

In this style, build an MDM centralized repository with curated master data, allows Data Steward to correct data in the MDM tool, cleansed master data are sync up with source systems.



**Key highlights:**

* Data consolidated into the MDM repository from multiple sources
* Updates to Master Data within MDM and source systems, continuous synchronization
* Data Stewards can clean master data into MDM
* Single version of truth between MDM and source systems

**Pros:**

* Cleanse master data are in sync between MDM and source systems
* Standard MDM Configuration

**Cons:**

* More intrusive as impacts source systems

**Possible Use Cases:**

* When multiple source systems and systems are able to take cleansed data automatically

## Transaction/Centralized

In this style, there is no source system to maintain master data and MDM tool becomes the system of records. This pattern is used mostly to manage reference data.



**Key highlights:**

* All Master Data is managed in MDM, sole provider of source of truth
* Master Data is authored in MDM, not in source systems
* All systems subscribe to the MDM repository for Master Data

**Pros:**

* MDM is the system of truth and source system

**Cons:**

* Most intrusive as MDM becomes source system
* Most MDM configuration

**Possible Use Cases:**

* When there are no source systems e.g. Reference data like Postal data

# The Standards relevant to MDM

This Guideline has been designed to provide tools in support of the following subset of Data Management Standard Requirements and Principles. The key Principle being: “Data has a single authoritative source”.

|  |
| --- |
| **PRINCIPLE 1 - Data is an Asset** |
| * **1.6** Establish and maintain a data inventory |

|  |
| --- |
| **PRINCIPLE 3 - Data is Fit for Purpose** |
| * **3.4** Who creates/modifies the data and in what process. * **3.6** Define what the authoritative source is and how it is provided for each area / set of data |

|  |
| --- |
| **PRINCIPLE 4 - Data is standardised** |
| * **4.4** Establish and maintain a change control process for data definitions and models * **4.5** Verify that any new solution designs align with agreed definitions & models |

|  |
| --- |
| **PRINCIPLE 5 - Data has a single authoritative source** |
| * **5.1** Establish and maintain a catalogue of all available data sources * **5.2** Define the single (master) trusted source of data where there are multiple copies of a dataset * **5.3** Ensure that any replicated data used for reporting or analysis is traceable to the master source[s] |

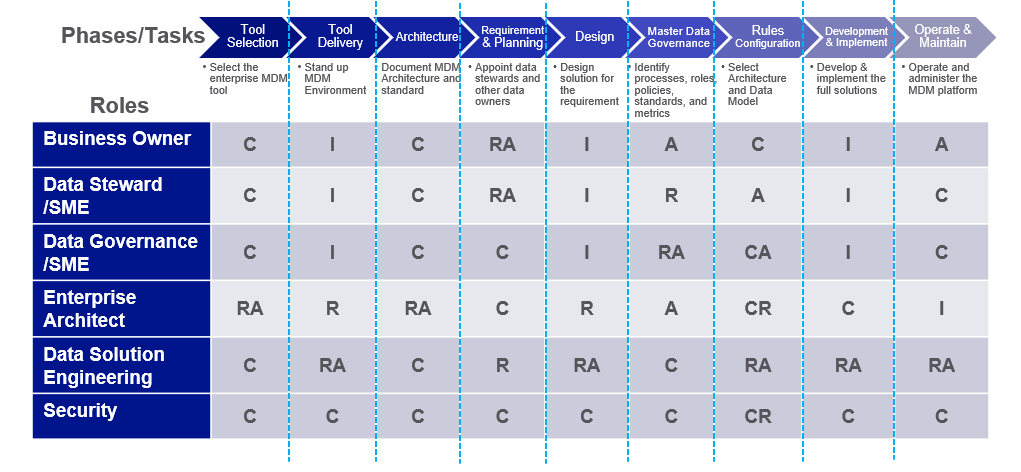
# MDM Best Practices

The following are the best practices of the Master Data Management:

* **MDM is not one and done:**  Master Data Management strategy should be woven into the foundation of our business. If data alignment is considered merely a one-time occurrence, we will encounter the same data mismanagement issues repeatedly.
* **Education is key:** All personnel and departments must be trained and regularly retrained on how to format, enter, store, and access data.
* **Start small, but think big:** When rolling out a new master data management strategy, we want to first focus on a smaller data set that may be causing some current business pain (e.g. customer or workforce or asset data for a specific geography). Done this way, you can assure buy-in for a larger rollout.
* **Keep our eye on ROI:** Since business units have different objectives, a common ROI should be established at the outset of a master data management strategy development, and the return on investment should be examined after each phase of rollout to maintain buy-in.
* **Buy-in at the top:** For master data management to be successful, leaders within all business units must be engaged in the development of the strategy, as well as continuously involved in ongoing governance conversations.
* **Don’t forget to update:** Master data management strategy must include regular, synchronized updates to ensure that our single source of data has the most accurate information.
* **Business ownership and governance:** Business “users” (stewards) must take full ownership of a master data in conjunction with the IT group for technology support. Also data governance is key to success of the master data management.
* **Change management:** Organizational change and knowledge transfer are the biggest master data challenges. A change management team, with a well thought out plan suitable for the client culture is critical. This team provides the leadership and fosters communication to resolve issues throughout the organization.
* **Make the Data Easy to Access:** Better organization means easier access. But, improving the retrieval process doesn’t stop there. Don’t put up so many barriers between your team and the data they need. This will only slow down the entire productivity of the business, and sometimes, it can confuse employees. Keep this from happening by making access easy for all.
* **Improve Data Security:** Both internal data and those purchased on the market must be protected to prevent information useful to the business from being compromised by a cyber-attack and security breach. Use user-specific data control to control data at the local level so that individual systems and users can receive only the data they need, only at the time they need it.

# MDM Roles and Responsibilities

The following is the RACI (**R**esponsible, **A**ccountable, **C**onsulted, **I**nformed) roles and responsibilities of the Master Data Management:



# MDM Architecture

The following diagram depicts the main areas of the Master Data Management architecture.



**The operational source systems** support the actual business processes and store the Master Data as well as the transaction data.

The Master Data Management platform

**The cloud data platform** is one of the main final recipients of the Master Data, generated and managed by the supporting components. The cloud data platform is the enterprise consolidated repository of both master and transactional data for the company.

**The business intelligence** reporting is performed using data from the cloud data platform data. The quality of the reporting generated is fully dependent on the quality of the Master Data.

**The third-party** vendordata like D&B, industry standard reference data is used to enrich the master data.

**The downstream applications/systems** can get the quality master data in real-time via API or in batch from MDM repository

The Data Solution delivery team handles the development, deployment and maintenance of the master data management platform. This area is not really in the scope of Master Data Architecture but is mentioned in this document at a high level for completeness.

These four components should satisfy both the differing user requirements as well as the ability to interact with one another seamlessly. This results in a balance that needs to be achieved between “best of breed” on the one hand and a high level of standardisation on the other hand.

By developing enterprise-wide standards for each of these areas and implementing them in a systematic step by step approach, National Grid could steadily converge to an application landscape which is more cost effective to roll out and maintain and which provides the business with a high degree of consistency in the operational and strategic information managed in those systems.

A Master Data architecture would provide National Grid with the tools and approach consisting of:

* An enterprise Master Data management system for enterprise-wide data standards
* A standard toolset for data governance, data management, design, Modeling, publishing and collaboration
* The first set of common data elements which will start to drive standardisation in reporting and in application configurations

This type of vision will not be achieved in a big bang approach in which all data, governance and application aspects are re-designed and re-implemented in one effort. The approach follows the following principles:

1. Whenever new applications and data work is performed, it should be ensured that the MDM work is future proof and scalable where appropriate. i.e. it fits into the global architecture and data model and is covered by the governance model
2. Ensuring that the overall architecture is initially designed as scalable doesn’t mean all components must immediately be built and would usually be done incrementally.
3. Develop the global standards in a top down approach, defining and standardising the most important elements first
4. Ensure that the architecture can cater for the integration of applications which are not (yet) compliant with the global standards. This is generally achieved by providing the tools to manage the necessary mappings for the non-compliant elements in those applications. This flexibility is key to the ability to quickly integrate acquisitions at least at a reporting level.

# Useful References

[Master Data Management Presentation](https://teams.microsoft.com/l/file/42B9B484-ACD7-46B1-973F-93490F7B47B9?tenantId=f98a6a53-25f3-4212-901c-c7787fcd3495&fileType=pptx&objectUrl=https%3A%2F%2Fnationalgridplc.sharepoint.com%2Fsites%2FGRP-INT-Architecture%2FShared%20Documents%2FData%20and%20Information%20Architecture%2FMaster%20Data%20Management%2FMaster%20Data%20Management%20draft.pptx&baseUrl=https%3A%2F%2Fnationalgridplc.sharepoint.com%2Fsites%2FGRP-INT-Architecture&serviceName=teams&threadId=19:0208706c699049ac805a806775adce1f@thread.skype&groupId=6e41877a-9b63-4354-b522-463045e111f7)

[Reltio MDM Product Demo](https://web.microsoftstream.com/video/7ed0caed-3147-479c-96c0-022610d09264)

# Document Administration

|  |  |  |
| --- | --- | --- |
| Owner: |  | Dan Robertson |
| Author/s first edition: |  | J Harrison |
| Author/s current edition: |  | Nishit Ajwaliya |
| Date first created: |  | 11th November 2016 |
| Date of current edition: |  | 9th August, 2021 |
| Date last amended: |  | 9th August, 2021 |

**Reviewers**

|  |  |  |
| --- | --- | --- |
| Review Name | Role | Date & version |
| Ben Howat | UK CoP – ETO | v1.1c-20th March 2018 |
| Paul Wilson | UK CoP - SO | v1.1c-20th March 2018 |
| David Harrison | UK CoP - GTO | v1.1c-20th March 2018 |
| Sean Dale | CoP - IS | v1.1c-20th March 2018 |
| Glen Truman | CoP - Legal | v1.1c-20th March 2018 |
| David Smith | UK CoP – HR | v1.1c-20th March 2018 |
| Wendy Arrowsmith | Global IS (DR&S) | v1.1c-20th March 2018 |
| Preston Large | US CoP – Data Management, Electric Operations | v1.1c-20th March 2018 |
| Nick Raad | US CoP – Gas Business Enablement and Gas Operations | v1.1c-20th March 2018 |
| Sue Sawchuk | US CoP – Bus Svc DM | v1.1c-20th March 2018 |
| Philip Kluko | US CoP – Bus Svc DM | v1.1c-20th March 2018 |
| Bill Donoghue | US CoP & lead for US Finance Master Data | v1.1c-20th March 2018 |
| Bonnie Burkhardt | US CoP – Bus Svc DM | v1.1c-20th March 2018 |
| Balaji Doraibabu | US - CoE | v1.1c-20th March 2018 |
| Sarah Gurr | UK - CoE | v1.1b-9th March 2018  v1.1c-20th March 2018 |
| John Purkins | UK - CoE | v1.1b-9th March 2018  v1.1c-20th March 2018 |
| Lisa Russell | UK - CoE | v1.1b-9th March 2018  v1.1c-20th March 2018 |
| Ben Morgan | BMS | v1.1c-20th March 2018 |

*CoP = Data Management Communities of Practice, CoE = Data Management Centre of Excellence*

**Approval**

|  |  |  |
| --- | --- | --- |
| Name | Role | Date |
| Dave Hopkinson | Head of Data Management CoE - UK | v1.0 - 25th April 2018  V2.0 – 30th Nov 2018 |
| Balaji Doraibabu | Director, Data Management CoE - US | v1.0 - 25th April 2018  V2.0 – 30th Nov 2018 |
| Dan Robertson | Head of Data and Analytics | V3.0 - |

**Change History**

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Author | Description of Changes |
| v2.0 | 30th Nov 2018 | Lisa Russell  Katie Brunton | Additions inserted from US comments |
| V3.0 | 9th Aug 2021 | Nishit Ajwaliya | Complete update of the document |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Superseded Documents**

|  |  |  |  |
| --- | --- | --- | --- |
| Reviewed by | Author(s) | Date & version | Description of changes |
| Sarah Gurr, Tracey Thompson | J Harrison | Nov 14th, 2016  V1.0 | Published version - DM-05-06 Data Reuse and Control - superseded by this document |
| Dave Hopkinson, Balaji Doraibabu | Katie Brunton / Lisa Russell | Nov 12th, 2018  V2.0 | Additions inserted from US comments |