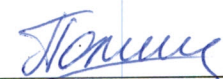


English title / Название на английском	DICTIONARY MANAGEMENT AND DATA CODING
Russian title / Название на русском	УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ
Number / Номер	DM010
Version / Версия	ver.2.0
Previous version / Предыдущая версия	/ № DataMatrix-SOP DM010/001 - October 2014
Effective Date ¹ / Дата вступления в силу	21. 08. 2015
Review Date / Дата пересмотра	21. 08. 2010

Prepared by	CDM	Tatiana Medvedeva
Подготовлено	Менеджер отдела ОКД	Татьяна Медведева

	10.08.2015
Подпись/Signature	Дата/Date

Reviewed by	QA Manager	Nadezhda Polischuk
Проверено	Менеджер по качеству	Надежда Полищук

	10.08.2015
Подпись/Signature	Дата/Date

Approved by	CEO	Kristina Leus
Утверждено	Генеральный директор	Кристина Леус

	10.08.2015
Подпись/Signature	Дата/Date

¹ The date format is Day/Month/Year.
Формат даты - День/Месяц/Год

Document History / Изменения в документе

Number & Version / Номер и версия	Date / Дата²	Changes	Изменения
№ DataMatrix-SOP DM010/001 - October 2014	October 30, 2014	New procedure	Новая процедура
DM010_ver.2.0	21. 08. 2015	SOP translated into Russian. Section 4.1. added. In Section 4.3. the process of dictionary version change described. In Section 5.1. description of types of coding added. Change of Dictionary Version Form in Russian and English created and approved.	Выполнен перевод на русский. Добавлен п.4.1. В п.4.3. описан процесс изменения версий словарей. В п.5.1. добавлено описание видов кодирования. Создана и утверждена форма «Изменение версии словаря».

² Effective date in format: Day/Month/Year.
Дата вступления в силу в формате: День/Месяц/Год

Contents / Содержание

1. Purpose, Scope and Key Aspects.....	4
1. Цель, сфера применения и общие положения	4
2. Applicable to	4
2. Применимо к	4
3. Terms and Definitions	4
3. Термины и определения	4
4. Dictionary Management.....	5
4. Управление словарями.....	5
5. Data Coding	7
5. Кодирование данных	7
6. References.....	10
6. Документы, упоминаемые в тексте	10
7. Appendices	10
7. Приложения	10

DICTIONARY MANAGEMENT AND DATA CODING	УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ
<p>1. Purpose, Scope and Key Aspects</p> <p>The purpose of this SOP is to describe the procedure of data coding and management of coding dictionaries used in Data [MATRIX] EDC. A variety of dictionaries may be used in the processing, analyzing and reporting of data collected in clinical trials. The most often used dictionaries are MedDRA (Medical Dictionary for Regulatory Activities) and WHO Drug (World Health Organization Drug Dictionary). The dictionaries used in particular studies should be declared in study-specific Data Management Plan (DMP) as well as other data coding specifications not covered in this SOP.</p> <p>2. Applicable to</p> <p>Head of Data Management Department (HDM), Senior Clinical Data Manager (SCDM), Clinical Data Manager (CDM), Data Entry Manager (DEM), Data Management Assistant (DMA), Database Administrator (DBA), Senior DBA.</p> <p>3. Terms and Definitions in English text</p> <p>AE – Adverse Event DMP – Data Management Plan EDC – Electronic Data Capture system MedDRA – Medical Dictionary for Regulatory Activities SF - Standard Form SOP – Standard Operating Procedure WHODRUG – World Health Organization Drug Dictionary</p>	<p>1. Цель, сфера применения и общие положения</p> <p>Данная процедура описывает процесс кодирования данных и управления словарями для кодирования используемых в системе Data [MATRIX] EDC. При обработке, анализе и предоставлении данных, собранных в процессе клинического исследования, могут быть использованы различные словари. Наиболее часто используются: MedDRA (Медицинский словарь терминологии регулятивной деятельности) и WHO Drug (Словарь Всемирной Организации Здравоохранения). Словари, применяемые в конкретных исследованиях, а так же иные спецификации по кодированию данных, не упоминаемые в данной СОП, следует указывать в проектно-м плане ОКД (DMP).</p> <p>2. Применимо к</p> <p>Руководитель отдела ОКД, старший менеджер отдела ОКД, менеджер отдела ОКД, менеджер ввода данных, ассистент отдела ОКД, администратор БД, старший администратор БД.</p> <p>3. Термины и определения в русском тексте</p> <p>EDC (Electronic Data Capture system) – электронная система ввода данных MedDRA (Medical Dictionary for Regulatory Activities) - медицинский словарь регулятивной деятельности. SF (Standard Form) – стандартная форма WHO Drug (World Health Organization Drug Dictionary) - словарь лекарственных средств Всемирной Организации Здравоохранения БД – база данных Интегрированные данные - данные об исследуемом препарате, собранные с разных фаз исследования. НЯ – нежелательное явление ОКД – обработка клинических данных СОП – стандартная операционная процедура</p>

DICTIONARY MANAGEMENT AND DATA CODING	УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ
<p data-bbox="136 228 515 263">4. Dictionary Management</p> <p data-bbox="136 263 492 298">4.1. Types of Dictionaries</p> <p data-bbox="136 298 1120 550">A variety of medical coding dictionaries may be used to process, analyze and report collected data. These dictionaries range in size and complexity from simple code lists with a few entries to large and complex dictionary systems containing thousands of entries and related tables. Two examples of commonly used dictionaries are the <i>Medical Dictionary for Regulatory Activities</i> (MedDRA) and the <i>World Health Organization Drug Dictionary</i> (WHO Drug).</p> <p data-bbox="136 694 1120 869">Custom dictionaries are typically developed to meet company-specific processes. Most custom dictionaries display terminology in a hierarchical pathway ranging from broad terminology to very specific terms. These dictionaries can be used to code adverse event data, medical history data and more commonly, medication data.</p> <p data-bbox="136 949 577 984">4.2. Dictionary Change Control</p> <p data-bbox="136 984 1120 1236">The selected dictionary and version should be documented in DMP, so that it can be referenced in clinical study reports or integrated summaries that report on the coded terms. The coding dictionaries should be securely maintained. If modifications are made in-house to a published reference, it should be clearly stated, so as not to mislead reviewers, who are familiar with the published reference. Any changes made to dictionary entries should be included in an audit trail.</p>	<p data-bbox="1120 228 1518 263">4. Управление словарями</p> <p data-bbox="1120 263 1422 298">4.1. Виды словарей</p> <p data-bbox="1120 298 2098 662">Разнообразные медицинские словари могут быть использованы для кодирования при обработке, анализе и представлении полученных данных. Эти словари различаются по размерам и сложности, начиная от простых списков для кодирования, содержащих небольшое количество словарных статей до больших и сложных словарных систем, содержащих тысячи словарных записей и связанных таблиц. Два наиболее широко используемых словаря: <i>Медицинский словарь регуляторной деятельности</i> (MedDRA) и Словарь лекарственных средств ВОЗ (WHO Drug).</p> <p data-bbox="1120 662 2098 917">Пользовательские словари обычно разрабатываются в соответствии с процессами конкретной компании. Большинство пользовательских словарей отображают терминологию в иерархической структуре от терминологии с широким смыслом до очень специфичных терминов. Эти словари могут использоваться для кодирования данных о нежелательных явлениях, данных медицинского анамнеза и, чаще всего, для данных о лекарственных средствах.</p> <p data-bbox="1120 949 1720 984">4.2. Управление изменениями словарей</p> <p data-bbox="1120 984 2098 1359">Выбранный словарь и его версия должны быть указаны в плане ОКД так, чтобы на документ можно было бы ссылаться в клинических отчетах или сводных обзорах, где упоминается кодирование терминов. Словари для кодирования должны храниться в надежном месте. Если в официально опубликованную версию вносятся модификации собственными силами компании, то следует четко прописывать изменения с тем, чтобы не запутать проверяющих лиц, которые знакомы с официально опубликованным вариантом. О любых изменениях в словарных статьях должен оставаться документальный след.</p>

DICTIONARY MANAGEMENT AND DATA CODING

4.3. Dictionary Version Control

The dictionary and version used for a given project, time period or sets of data should be clearly documented in Data Management Plan (DMP). When there are multiple ongoing studies, a determination should be made of which dictionary and version will be used for the coding for each study and a systematic process and instructions should be in place to ensure the use of the appropriate dictionary and version. The evaluation of the extent of the changes between versions, their impact on previously coded terms and the criteria for implementing the latest version and recoding should be established in DMP.

The dictionary and version information can be maintained within the clinical database or within the metadata of data sets containing coded data.

In case of dictionary version change, CDM/SCDM should complete the corresponding form (see *SF – Change of Dictionary Version*) and approve changes with sponsor representative and/or project manager. Then the version change request should be sent to the responsible DBA. When request is realized, DBA should complete the form (sign and date) and give it back to Data Management department. The completed form is kept in the project Data Management File.

4.4. Reporting of Coded Terms

All levels, codes or group assignments for the data should be stored. The risk of not storing all coding information is the necessity to recode or retrieve the additional coding information for reports at other levels or groups that are needed at a later date.

УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ

4.3. Контроль версий словарей

Словарь и его версия, используемые для конкретного проекта, промежуток времени или наборы данных следует четко описывать в плане ОКД. При наличии нескольких текущих исследований необходимо определить, какой словарь и какая версия будут использованы для кодирования в каждом исследовании, а для обеспечения использования подходящего словаря и версии, необходимо проводить систематическое инструктирование. Оценку изменений в разных версиях, их влияние на ранее закодированные термины, критерии для внедрения последней версии словаря, а также причины перекодирования данных следует описывать в плане ОКД.

Информацию о словаре и его версии можно указывать в клинической БД или в метаданных наборов данных, содержащих закодированные данные.

При смене версии словаря менеджеру отдела ОКД следует заполнить соответствующую форму (см. *SF – Изменение версии словаря*) и утвердить изменения у представителя спонсора и/или проектного менеджера. Далее запрос на изменения версии словаря направляется администратору БД. После выполнения запроса администратор БД заполняет форму (ставит дату выполнения и подпись) и передает ее в отдел ОКД. Заполненную форму следует хранить в проектной папке ОКД.

4.4. Отчеты о закодированных терминах

Следует сохранять все уровни, коды или распределение данных по группам. Невозможность сохранить всю закодированную информацию приводит к необходимости перекодирования или получения дополнительной информации о кодируемом термине для отчетов на разных уровнях или групп, необходимых на более

DICTIONARY MANAGEMENT AND DATA CODING

Changes to the dictionary and version may occur during the course of a project across individual studies. Ideally, the integrated data that is combined should be coded with a single dictionary version. However, an impact analysis between versions may help to identify the necessity for recoding data for integrated studies. If deemed necessary to recode data, it would only be required for those studies that were coded with a different version. Recoding data minimizes the resources and time needed to prepare the integrated data, while allowing for the utilization of the latest version for analysis. In most cases, the evaluation of safety is performed at the integrated level, rather than at the individual study report level. If uniform coding is not employed, the use of different dictionaries or versions among the individual study reports and integrated summaries should be documented.

5. Data Coding**5.1. Types of Coding**

Auto-encoding is a programmatically assisted process for matching a reported term to a dictionary term. Various methods exist for auto-encoding, such as character string matches with the dictionary, character string matches with synonym lists, and matches found using algorithms.

Manual coding refers to a situation where a person selects an appropriate dictionary entry for each reported term, either in the patient database or in a module of the dictionary application that deals with discrepancies. This method may be used when an auto-encoder is unable to code a term or an auto-encoder is not being used.

УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ

позднем этапе.

Изменения в словаре и версии могут возникать в отдельных исследованиях в ходе проекта по разработке лекарственного препарата. В идеале интегрированные данные должны быть закодированы с помощью одной версии словаря. Однако анализ изменений в разных версиях помогает определить необходимость перекодирования данных в комплексных исследованиях. Считается, что необходимо перекодировать данные только в тех исследованиях, где термины были закодированы с разными версиями словарей. Перекодирование данных сокращает ресурсы и время, необходимые для подготовки интегрированных данных, при этом позволяя использовать последнюю версию для анализа. В большинстве случаев оценка безопасности выполняется на интеграционном уровне, а не на уровне создания отчета для конкретного исследования. Если не используется единая кодировка, использование разных словарей или их версий должно быть задокументировано в отчетах для конкретного исследования и сводных обзорах.

5. Кодирование данных**5.1. Виды кодирования**

Автоматическое кодирование – запрограммированный процесс согласования указанного термина со словарным. Существуют различные методы автоматического кодирования, такие как: соответствие символьной строки со словарной, согласование символьной строки со списками синонимов, а так же соответствия, найденные с помощью алгоритмов.

Ручное кодирование соответствует ситуации, когда человек выбирает соответствующую словарную статью для каждого термина, полученного либо из базы данных о пациентах, либо из модуля словарного приложения, который обрабатывает несоответствия. Данный метод может быть применим, когда невозможно

DICTIONARY MANAGEMENT AND DATA CODING	УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ
<p>A hybrid approach to coding uses an auto-encoder to first automatically code those reported terms that match a dictionary term or that match a term that has previously been coded (i.e., a synonym list). The terms that are not auto-encoded are then manually coded.</p> <p>Auto-encoding is highly recommended to executing a dictionary against AEs. Training of the monitors and site personnel to cultivate an understanding of the coding process should facilitate capture of AE data in a format that can be auto-encoded. Training should include guidelines such as the following:</p> <ul style="list-style-type: none"> – Avoid use of adjectives as initial words (e.g., “weeping wound” may be coded to “crying” or “faint rash” may be coded to “syncope”) – Avoid use of symbols and abbreviations in the AE text field (symbols may be interpreted by different individuals differently) – Avoid inclusion of severity (e.g., “severe headache” in the AE text box inhibits auto-encoding; severity should be recorded in the severity box, not the AE text box) – Ensure AE text has clinical meaning (e.g., “bouncing off the walls” and “feeling weird” are difficult to interpret) – Ensure AE text has clear meaning (e.g., “cold feeling” may be interpreted as “chills” or “flu symptoms”). 	<p>закодировать термин с помощью автоматического кодирования или автоматическое кодирование не используется.</p> <p>Полуавтоматическое кодирование или гибридный подход к кодированию подразумевает первичное применение автоматического кодирования для тех терминов, которые соответствуют словарному термину или термину, который был ранее закодирован (напр., список синонимов). Термины, которые не были закодированы автоматически, затем кодируются вручную.</p> <p>Рекомендуется использовать автоматическое кодирование НЯ. Обучение мониторов и сотрудников исследовательских центров, объясняющее процесс кодирования, должно облегчить сбор данных НЯ в формате, подходящем для автоматического кодирования. Обучение должно основываться на следующих принципах:</p> <ul style="list-style-type: none"> – Избегайте использования прилагательных в начале словосочетаний (напр., «мокнущая рана» может быть закодирована как «мокрая экзема» или «мокрый дерматит»). – Избегайте использования символов и аббревиатур в текстовом поле НЯ (разные люди могут по-разному интерпретировать разные символы). – Избегайте указаний на тяжесть состояния (напр., «сильная головная боль» в текстовом поле НЯ препятствует корректному автоматическому кодированию; тяжесть состояния следует указывать в соответствующем поле, а не в текстовом поле НЯ). – Убедитесь, что в тексте НЯ имеется описание клинически значимого события (напр., с трудом поддаются интерпретации такие описания как «падение» или «плохое самочувствие») – Убедитесь, что текст описывает конкретное НЯ (напр., «пациенту холодно» может быть истолковано как «озноб», «переохлаждение» или «симптом гриппа»)

DICTIONARY MANAGEMENT AND DATA CODING

Encoding within the database requires the final medical judgment. After database lock, if medical judgment indicates that the default pathway inaccurately captures the medical condition, a change would require unlocking the database if the auto-encoding is done within the database proper. This practice does impose the need for an audit trail on analysis files.

5.2. Hard-Coding

Hard-coding or coding outside the clinical database is sometimes used to introduce medical judgment that the standard dictionary does not offer. When events such as “strange feeling” are reported and no additional information from the site is available, the medical monitor for the study may have insight that assists with the codification of the event, which can be inserted into the AE analysis file through hard-coding. It is possible to use “pass-through” text for the AE preferred term (conventionally, many sponsors make use of quotation marks to indicate verbatim text that is passed through by a program to the preferred term field) and hard-code the body system. Any use of hard-coding requires careful documentation.

5.3. Data Manager’s Impact in Data Coding

When coding, SDCM/CDM or assigned employee should follow the instructions on coding data (see *WP – Data Coding Instructions*)

In case AE text includes more than one event, SCDM/CDM

УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ

Кодирование в базе данных требует итоговой оценки с медицинской точки зрения. После закрытия БД, если медицинское заключение указывает на то, что выбранный по умолчанию словарный термин неточно отражает медицинское состояние, его изменение потребует разблокирования БД, если автоматическое кодирование производится непосредственно в БД. В подобных случаях возникает необходимость сохранять документальный след в анализируемых файлах.

5.2. «Жесткое» кодирование

Иногда для указания оценки с медицинской точки зрения, которую нельзя внедрить в стандартный словарь, используется «жесткое» кодирование или кодирование вне клинической БД. Когда сообщается о таких событиях, как «чувствует себя странно», а дополнительная информация из центра отсутствует, может помочь медицинский монитор, объяснение которого можно добавить в файл анализа НЯ в процессе «жесткого» кодирования. Возможно использовать сквозной термин в качестве предпочтительного термина НЯ (обычно многие спонсоры предлагают заключать определение термина в кавычки, которые программа пропускает и переходит непосредственно к предпочтительному термину) и закодировать систему органов тела, к которой относится термин. Любое использование «жесткого» кодирования требует тщательного документирования.

5.3. Влияние менеджеров по ОКД на кодирование данных

При кодировании данных менеджер отдела ОКД или назначенный сотрудник руководствуется инструкциями по кодированию данных (см. *WP - Инструкции по кодированию данных*).

В случае если текст описания НЯ содержит более одного

DICTIONARY MANAGEMENT AND DATA CODING

responsible for data coding should query the site in order to split the terms. When two events are reported in the same text field (e.g., “indigestion and diarrhea”), if splitting is done by the data management staff rather than the site, inconsistencies within the database may result. It’s not allowed to use splitting or lumping during data coding. Any terms modifications should be requested to the site/investigator.

6. References

WP DM016 – Data Coding Instructions

7. Appendices

Type	Code	Title	Language
SF	DM026	Change of Dictionary Version	Eng

УПРАВЛЕНИЕ СЛОВАРЯМИ И КОДИРОВАНИЕ ДАННЫХ

события, для разделения терминов менеджеру отдела ОКД, ответственному за кодирование данных, следует направить запрос на уточнение данных в исследовательский центр. Когда два события указаны в одном поле (напр., «несварение желудка и диарея»), а разделение терминов осуществляется сотрудниками отдела ОКД, а не исследовательским центром, то это может привести к несоответствиям в БД. Поэтому не допускается разделять термины в процессе кодирования данных. На любые изменения в терминах следует направлять запрос исследователю или в центр.

6. Документы, упоминаемые в тексте

WP DM016 - Инструкции по кодированию данных

7. Приложения

Тип	Номер	Название	Язык
SF	DM026	Изменение версии словаря	Ru