EDF Header Validation Requirements

# 1. Introduction

# This document provides detailed description of the validation requirements for EDF files. EDF-Editor-Translator[1] tool will provide “Validation Tool” for examining every EDF Header both in single-file mode and batch mode.

# The tool will

# Provide a detailed log of EDF files that do not comply to the EDF Header specification with description of the header parameters that failed.

# The tool will provide an interface to automatically fix the errors that were reported in the EDF Header. After verification by the user, the EDF file header can be permanently changed to resolve the errors

# 2. Background

## 2.1 EDF Specification

As defined in EDF, below is the detailed digital format of the header record (upper block, ascii's only) and of each subsequent data record (lower block, integers only). Note that each one of the ns signals is characterized separately in the header.

HEADER RECORD (we suggest to also adopt the 12 simple additional EDF+ specs)

8 ascii : version of this data format (0)

80 ascii : local patient identification (mind item 3 of the additional EDF+ specs)

80 ascii : local recording identification (mind item 4 of the additional EDF+ specs)

8 ascii : startdate of recording (dd.mm.yy) (mind item 2 of the additional EDF+ specs)

8 ascii : starttime of recording (hh.mm.ss)

8 ascii : number of bytes in header record

44 ascii : reserved

8 ascii : number of data records (-1 if unknown, obey item 10 of the additional EDF+ specs)

8 ascii : duration of a data record, in seconds

4 ascii : number of signals (ns) in data record

ns \* 16 ascii : ns \* label (e.g. EEG Fpz-Cz or Body temp) (mind item 9 of the additional EDF+ specs)

ns \* 80 ascii : ns \* transducer type (e.g. AgAgCl electrode)

ns \* 8 ascii : ns \* physical dimension (e.g. uV or degreeC)

ns \* 8 ascii : ns \* physical minimum (e.g. -500 or 34)

ns \* 8 ascii : ns \* physical maximum (e.g. 500 or 40)

ns \* 8 ascii : ns \* digital minimum (e.g. -2048)

ns \* 8 ascii : ns \* digital maximum (e.g. 2047)

ns \* 80 ascii : ns \* prefiltering (e.g. HP:0.1Hz LP:75Hz)

ns \* 8 ascii : ns \* nr of samples in each data record

ns \* 32 ascii : ns \* reserved

## 2.2 EDF+ Specifications

1. In the header, use only printable US-ASCII characters with byte values 32..126.
2. The 'startdate' and 'starttime' fields in the header should contain only characters 0-9, and the period (.) as a separator, for example "02.08.51".
3. In the 'startdate', use 1985 as a clipping date in order to avoid the Y2K problem. So, the years 1985-1999 must be represented by yy=85-99 and the years 2000-2084 by yy=00-84. After 2084, yy must be 'yy' and only item 4 of this paragraph defines the date.
4. The 'local patient identification' field must start with the subfields (subfields do not contain, but are separated by, spaces):
   1. The code by which the patient is known in the hospital administration.
   2. Sex (English, so F or M).
   3. Birthdate in dd-MMM-yyyy format using the English 3-character abbreviations of the month in capitals. 02-AUG-1951 is OK, while 2-AUG-1951 is not.
   4. Patients name.
5. Any space inside the hospital code or the name of the patient must be replaced by a different character, for instance an underscore. For instance, the 'local patient identification' field could start with: MCH-0234567 F 02-MAY-1951 Haagse\_Harry. Subfields whose contents are unknown, not applicable or must be made anonymous are replaced by a single character 'X'. Additional subfields may follow the ones described here.
6. The 'local recording identification' field must start with the subfields (subfields do not contain, but are separated by, spaces):
   1. The text 'Startdate'.
   2. The startdate itself in dd-MMM-yyyy format using the English 3-character abbreviations of the month in capitals.
   3. The hospital administration code of the investigation, i.e. EEG number or PSG number.
   4. A code specifying the responsible investigator or technician.
   5. A code specifying the used equipment.
7. Any space inside any of these codes must be replaced by a different character, for instance an underscore. The 'local recording identification' field could contain: Startdate 02-MAR-2002 PSG-1234/2002 NN Telemetry03. Subfields whose contents are unknown, not applicable or must be made anonymous are replaced by a single character 'X'. So, if everything is unknown then the 'local recording identification' field would start with: Startdate X X X X. Additional subfields may follow the ones described here.
8. 'Digital maximum' must be larger than 'Digital minimum'. In case of a negative amplifier gain the corresponding 'Physical maximum' is smaller than the 'Physical minimum'. Check item 9 on how to apply the 'negativity upward' rule in Clinical Neurophysiology to the physical ordinary signal. 'Physical maximum' must differ from 'Physical minimum'. In case of uncalibrated signals, physical dimension is left empty (that is 8 spaces), while 'Physical maximum' and 'Physical minimum' must still contain different values (this is to avoid 'division by 0' errors by some viewers).
9. Never use any digit grouping symbol in numbers. Never use a comma "," for a for a decimal separator. When a decimal separator is required, use a dot (".").
10. The ordinary signal samples (2-byte two's complement integers) must be stored in 'little-endian' format, that is the least significant byte first. This is the default format in PC applications.
11. The 'starttime' should be local time at the patients location when the recording was started.
12. Use the standard texts and polarity rules at<http://www.edfplus.info/specs/edftexts.html>. These standard texts may in the future be extended with further texts, a.o. for Sleep scorings, ENG and various evoked potentials.
13. The 'number of data records' can only be -1 during recording. As soon as the file is closed, the correct number is known and must be entered.
14. If filters (such as HighPass, LowPass or Notch) were applied to the ordinary signals then, preferably automatically, specify them like "HP:0.1Hz LP:75Hz N:50Hz" in the "prefiltering" field of the header. If the file contains an analysis result, the prefiltering field should mention the relevant analysis parameters.
15. The "transducertype" field should specify the applied sensor, such as "AgAgCl electrode" or "thermistor".

# 3. EDF Editor Header Validation Requirements

The EDF Header shall be validated for the following rules. All attributes that do not comply will be logged in the error log.

The user/verifier can visually examine the EDF file for these errors and provide corrective measures for the data in the header.

|  |  |
| --- | --- |
| **Description in EDF** | **Validation Rules Examined by Editor** |
| version of this data format (0) | (A.1) check for ascii  (A.2) is equal to 0  (A.3) cannot be empty field |
| local patient identification | (B.1) check for ascii  (B.2) can be empty field |
| local recording identification | (C.1) check for ascii  (C.2) can be empty field |
| startdate of recording (dd.mm.yy) | (D.1) check for ascii  (D.2) cannot be empty field  (D.3) dd:00-31, mm:00-12, yy:00-99  (D.4) separator between digits should be only ‘period’ |
| starttime of recording (hh.mm.ss) | (E.1) check for ascii  (E.2) cannot be empty field  (E.3) hh:00-23, mm:00-59, ss:00-59  (E.4) separator between digits should be only ‘period’ |
| number of bytes in header record | (F.1) check for ascii  (F.2) should not be empty  (F.3) is a non-negative integer |
| reserved | (G.1) check for ascii  (G.2) can be empty field |
| number of data records (-1 if unknown) | (H.1) check for ascii  (H.2) should not be empty  (H.3) is a positive integer (-1 if unknown) |
| duration of a data record, in seconds | (I.1) check for ascii  (I.2) should not be empty field  (I.3) is a positive floating point number (eg. 1, 0.2, 0.001) |
| number of signals (ns) in data record | (J.1) check for ascii  (J.2) cannot be empty  (J.3) is a positive integer |
| ns \* label (e.g. EEG Fpz-Cz or Body temp) | (K.1) check for ascii  (K.2) no duplicate signal labels (there cannot be 2 signals with same label EKG; EKG and ekg are considered as duplicates)  (K.3) cannot be empty field |
| ns \* transducer type (e.g. AgAgCl electrode) | (L.1) check for ascii  (L.2) can be empty field |
| ns \* physical dimension (e.g. uV or degreeC) | (M.1) check for ascii  (M.2) can be empty field - (uncalibrated) |
| ns \* physical minimum (e.g. -500 or 34) | (N.1) check for ascii  (N.2) is a floating point number  (N.3) Postive amplitude gain: physical minimum < physical maximum and digital min < digital max  (N.4) Negative amplitude gain: physical min > physical max and digital min < digital max (NEEDS DISCUSSION)  (N.5) cannot be empty field (even if uncalibrated - to avoid divide-by-0) |
| ns \* physical maximum (e.g. 500 or 40) | (O.1) check for ascii  (O.2) is a floating point number  (O.3) physical maximum NOT = physical minimum  (O.4) cannot be empty field (even if uncalibrated - to avoid divide-by-0) |
| ns \* digital minimum (e.g. -2048) | (P.1) check for ascii  (P.2) is an integer  (P.3) since each date sample is a 2-byte integer, range check [-32768,32767]  (P.4) cannot be empty field |
| ns \* digital maximum (e.g. 2047) | (Q.1) check for ascii  (Q.2) is an integer  (Q.3) since each date sample is a 2-byte integer, range check [-32768,32767]  (Q.4) cannot be empty field  (Q.5) digital minimum < digital maximum  (Q.6) digital minimum NOT = digital maximum (division-by-0 condition) |
| ns \* prefiltering (e.g. HP:0.1Hz LP:75Hz) | (R.1) check for ascii  (R.2) can be empty field |
| ns \* nr of samples in each data record | (S.1) check for ascii  (S.2) is an integer  (S.3) is greater than 0  (S.4) cannot be empty fieldcd in |
| ns \* 32 ascii : ns \* reserved | (T.1) check for ascii  (T.2) can be empty field |

# 4. References

1. EDF-Editor-Translator. URL:<https://github.com/nsrr/edf-editor-translator>.

2. European Data Format. URL:<http://www.edfplus.info/specs/edf.html>.

3. European Data Format Plus. URL:<http://www.edfplus.info/specs/edfplus.html>.

4. Dennis Dean’s Matlab code work: [https://github.com/DennisDean/lockEdfLoadClass/blob/master/BlockEdfLoadClass.m](https://github.com/DennisDean/BlockEdfLoadClass/blob/master/BlockEdfLoadClass.m).