Dental Care Data
in the Health Information System

Tervise Arengu Instituut
National Institute for Health Development

National Institute for Health Development

Dental Care Data in the Health Information System

Hedi Liivlaid

Tallinn 2019

The **mission** of the National Institute for Health Development is to create and share knowledge for influencing the attitudes, behaviour, policies and the environment with evidence-based information with an aim of improving the well-being of the people in Estonia.

Contents

Definitions

Data field	a single variable of a data set
Dental formula	number of the tooth on which the respective diagnosis was given or the dental work performed
Dental record	a document describing the treatment, treatment plan and general anamnesis of the oral cavity of the patient. A unique dental record is a combination of a document number and a health care provider's business registry code. A dental record may contain several visits and one visit may involve several dental treatments performed
DMFT index	index on cariology of permanent teeth. Determines as the total of decayed (D), filled (F) and missing (M) teeth (T)
dmft index	index on cariology of deciduous teeth. Determines as the total of decayed (d), filled (f) and missing (m) teeth (t)
Document	in the context of this paper, the term "document" is used as a synonym for dental record, see "dental record"
Entry	a row of the data set to be analysed
Gingivitis	or inflammation of the gums is a reversible damage of the gum tissue

Abbreviations

EHIF	Estonian Health Insurance Fund
HCP	health care provider
HIS	Health Information System
ICD-10	International Statistical Classification of Diseases and Related Health Problems 10th Revision (https://icd.who.int/browse10/2016/en).
MoSA	Ministry of Social Affairs
NIHD	National Institute for Health Development
TEHIK	Tervise ja Heaolu Infosüsteemide Keskus (Health and Welfare Information Systems Centre)

Summary

In 2016, **61%** of health care providers forwarded dental care data to Health Information System in comparison with the sum what they sent to the National Institute for Health Development with their health care statistics reports. Such data coverage is not sufficient for compiling reliable statistics. It should also be kept in mind that some health care providers had sent a very small number of dental records to the Health Information System, thereby performing their obligation to provide data, but nevertheless, the volume of data sent is equally important.

Depending on the methodology for summarising the number of dental visits, the visits uploaded to the Health Information System made up **51–64%** of visits reported to the National Institute for Health Development.

In the part on dental care diagnoses, **one and two-digit** ICD-10 codes are used, what does not enable to produce diagnose based statistics.

The statistical category of diagnoses is often stated incorrectly – in 99% of the cases, the statistical type of the main diagnosis was marked as the primary one.

Health care providers use a variety of alternative lists for registering the work done, making unified statistics difficult to compile. In these instances, verbal entries in text fields should be used when analysing the work done for statistical purposes. With such a large amount of data, working with text fields is highly complicated, time-consuming and non-optimal.

Several important data fields in terms of statistics are left empty or only filled in by some of the health care providers. For example, indexes on cariology of teeth and the dental formula of the treatments performed were completely left out by all health care providers, and only one health care provider had filled in the gingivitis data block.

It is important to improve the coverage and quality of data before starting to produce dental health statistics on the basis of the Health Information System.

Introduction

Good health policy is based on high-quality health statistics. The availability of up-to-date, reliable, objective, and regular public health and health care system information is essential for making practical management decisions. At the same time, it is important to maintain an optimal reporting burden of health care providers (HCP).

One of the priorities of the National Institute for Health Development (NIHD) is to reduce the work load of the HCPs. NIHD finds that if the HCPs are already submitting data somewhere, it is unreasonable to gather such data several times and solutions should be developed for using the existing data.

In order to produce high-quality statistics and reduce the reporting burden, NIHD will continue to seek for possibilities of utilising alternative data sources. Health Information System (HIS)[1] is definitely one of the largest databases containing data necessary for health statistics. Thus it is all the more logical to develop health statistics by starting using a system which allows the use of already collected detailed data for producing more accurate and high-quality health statistics. NIHD has worked in that direction for years, carrying out regular analyses of health data received by the HIS and cooperating with professional organisations, software developers, the Ministry of Social Affairs (MoSA), Health and Welfare Information Systems Centre (TEHIK), and HCPs.

On 1 July 2015, dentists entered into an obligation to provide electronic dental records for HIS. Dental records contain information on dental services provided to patients. Dental records are based on the needs of dentists and, in theory, should also largely meet the needs for dental care statistics. Currently, dental statistics are produced on the basis of aggregated data collected by NIHD on a reporting form established by an order of the Minister responsible for the field.

Dental care statistics are primarily used within the country. The main users are MoSA, the Estonian Health Insurance Fund (EHIF) and professional associations in the field of health, but also the HCPs, journalists, students and pharmaceutical companies. The MoSA needs statistics for work in organising and planning the provision of health care services and for getting an overview of the service volumes provided; the EHIF uses the statistics published by the NIHD for planning the payment of dental care benefits, for instance. The indicator used in international comparisons is the number of visits per person (Eurostat/WHO/OECD Joint Questionnaire). At the same time, there is a demand for even more detailed statistics that the NIHD is currently unable to provide. For example, OECD requires for oral health index of 12-year-old children (DMFT index) as a part of its annual database inquiry.

Consumer expectations to associated information have increased over time. For example, information is needed on patients of different ages, both in dentistry and in orthodontics, by diagnoses and procedures performed. Making the aggregated reports more detailed would make submission of the required information too costly for the reporters. It is reasonable to compile more detailed statistics based on individual data. That is why the hopes for obtaining more detailed data are mainly placed on the information on dental records submitted to the HIS. If the HIS data is of high quality and the data is sufficient, dental care providers could be relieved from the obligation to submit dental care aggregated reports to the NIHD.

---

[1] The health information system is a data set in the national information system which processes data related to the area of health care for the conclusion and performance of agreements for the provision of health care services, for the assurance of the quality of health care services and of patients' rights and for the protection of public health, including for the maintenance of registers reflecting health status and for health care management (Health Services Organisation Act, § 59[1] (1)).

The report provides an overview of the amount and quality of dental care data provided to the HIS in 2016. The 2016 data gives an opportunity for the first time to analyse the annual data on dental care received by the HIS and compare them with data from aggregated dental visits collected by the NIHD.

The Department of Health Statistics at the National Institute for Health Development uses the results of the analysis for planning national health statistics. The results will also be sent to the Estonian Dental Association and to specialists at TEHIK and MoSA to plan the development of HIS and its bulk data extraction and to improve the data quality by reducing the occurrence of problems found in the future.

1 Initial data

The analysis is based on an extract from the HIS operational database of dental records received in 2016. Generally, the Department of Health Statistics at the National Institute for Health Development receives regular extracts from Health Information System data 30 days after the end of the quarter about the cases terminated in the previous quarter. Since dental care data has not yet been downloaded to the Statistics module, the extract had to be made from the HIS operational database. For that, a permission was applied from the MoSA and an extract made of the respective data from TEHIK. The data needed for the national dental care statistics was requested for analysis.

The data used in the analysis was as follows:

1. Document data
   1.1 Document number
   1.2 Document validation time
   1.3 Time of receipt of the document at HIS

2. Data of the document's compiler
   2.1. Number of health care professional's certificate of registration
   2.2. Name of the health care provider
   2.3. Commercial registry code of the health care provider

3. Patient information
   3.1. Patient's unique ID
   3.2. Sex
   3.3. Age
   3.4. Real residence

4. Dental care visit data
   4.1. Visit date
   4.2. Dental care diagnosis
    4.2.1. Main disease
       4.2.1.1. Code and name of the diagnosis according to ICD-10
       4.2.1.2. Verbal clinical diagnosis
       4.2.1.3. Statistical type of diagnosis
       4.2.1.4. Dental formula
    4.2.2. Complications from the main disease
       4.2.2.1. Code and name of the diagnosis according to ICD-10

   4.2.2.2. Verbal clinical diagnosis
   4.2.2.3. Dental formula
  4.2.3. Concomitant disease
     4.2.3.1. Code and name of the diagnosis according to ICD-10
     4.2.3.2. Verbal clinical diagnosis
     4.2.3.3. Statistical type of diagnosis
     4.2.3.4. Dental formula
  4.3. Data on works performed
   4.3.1. Code and name of the work performed
   4.3.2. Dental formula

5. Data on clinical examination in dental care
   5.1. Cariologic tooth indexes
      5.1.1. DMFT index value
      5.1.2. dmft index value
   5.2. Periodontological diagnosis of teeth
      5.2.1. Periodontological status code
      5.2.2. Periodontological status name
   5.3. Gingivitis data
      5.3.1. Code and name of the diagnosis
      5.3.2. Verbal clinical diagnosis
      5.3.2. Gingivitis level
      5.3.3. Dental formula

6. Data on summary of patient's treatment
   6.1. Reason for ending treatment
   6.2. Performance of sanitation

The number of HCPs providing dental care services and the total data submitted by them to the NIHD was compared with the HIS data. The visits submitted to NIHD are taken from the 2016 quarterly report "Arsti vastuvõtud ja koduvisiidid" ("Physicians' outpatient visits and home visits")[2] gathered from the HCPs, which includes visits to all dental professions. More detailed information on dental care services are gathered in an annual report "Hambaarsti aruanne" ("Dentist's report").

---

[2] Collection of quarterly data was stopped in 2017. From now on, dental visits will only be gathered in the report „Hambaarsti aruanne" ("Dentist's report").

2 Data quality

2.1 Data coverage

In 2016, 506 HCPs providing dental care services submitted their dental visits to the NIHD in the quarterly report "Physicians' outpatient visits and home visits". These HCPs had registered a total of 1,883,225 dental visits. With the "Dentist's report" to the NIHD, 505 HCPs stated that they provide dental care services. The number of HCPs who submitted a dentist's report was one less, because one institution joined the other one in the middle of a reporting year and the data of the dental care was presented jointly. Dental visits in the Health Information System are discussed in more detail in Chapter 2.2.4.

61% (324 HCPs) of the HCPs providing dental care data to the NIHD had submitted at least one dental record to HIS in 2016. By counties, the ones with the lowest participation rate are Lääne-Viru and Võru countries (Figure 1) where respectively 29% and 36% of the health care providers had sent dental records. The highest share of HCPs submitting dental care data to the HIS was in Jõgeva and Valga counties.
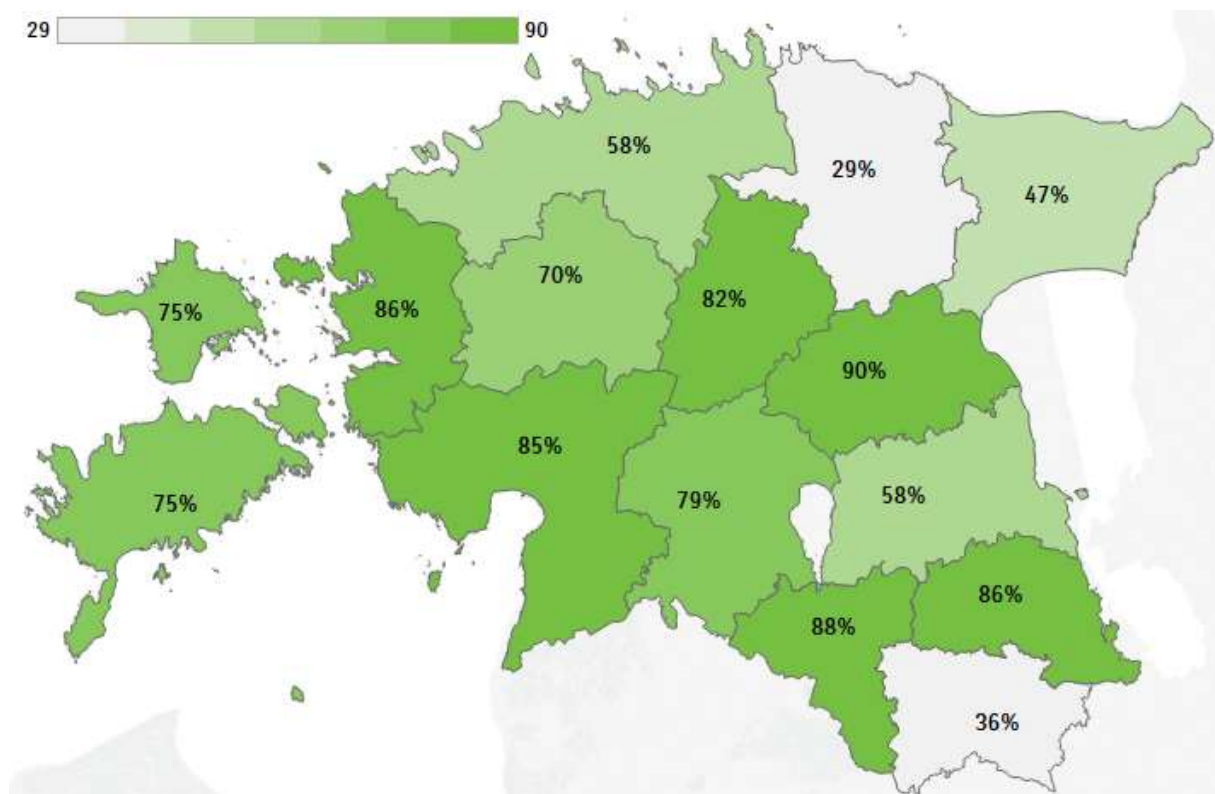

Figure 1. The share of health care providers who sent dental records to the Health Information System by county, 2016

17 HCPs provided less than ten dental records (Figure 2), but had a total of up to 34,000 visits according to the NIHD report. The maximum number of dental records presented by a HCP was 55,724. In total, 834,639 dental records were sent to the HIS.
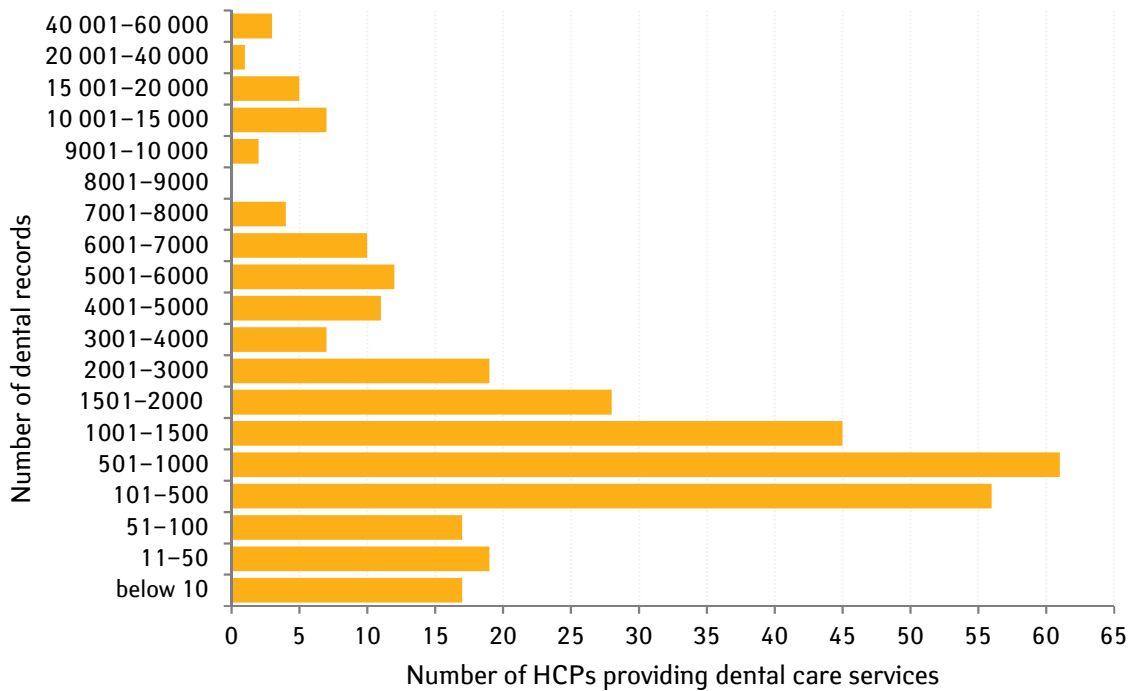
Figure 2. Number of dental records sent to Health Information System, 2016

There were also HCPs that did not send dental records to the HIS, but had indicated professional work of a dentist on outpatient epicrisis. Some of them were very large service providers whose work makes up an important part in the dental care statistics. For example, a one large hospital has not provided any dental records to the HIS, but has provided dentist's work data on outpatient epicrisis. Compared to the dentist's visits at this large hospital submitted to the NIHD, and visits submitted on an outpatient epicrisis, the data matches almost 100%. Dental care data provided on outpatient epicrisis was not included in this analysis.

2.2 Verification of the variables' fulfilment submitted to the HIS

Below is a description of HIS data variables and their fulfilment one by one.

2.2.1 Document data

Document number

The document number is generated by the HCPs information system in accordance with the rules for document numbering of each HCP. A document code of a unique HCP document is made up of the HCPs commercial registry code and document number because documents with the same number are sent by different HCPs.

Analyses of the data field data shows that different rules were used to fill in the field. It appeared that users of the software program Hammas developed by Innovaatik OÜ had a prefix in front of the document number, e.g., *inno.hambaravi.20160915143203343127359717.* For the point of view of compilation of statistics, such numbering system is not an obstacle.

Generated in the HCP's information system at the moment the document is signed.

Time of receipt of the document at HIS

Automatically saved in the Health Information System upon receipt of the document.

Since all fields are mandatory in the Health Information System's dental record data block and are automatically generated, there were no problems with filling of these fields that would prevent production of statistics.

2.2.2 Data of the document's compiler

Number of health care professional's certificate of registration

The health care professional's registration code, according to the Estonian Health Care Professionals Register, is entered by the health care professional or is automatically added on the basis of the information available in the HCPs information system. The analysis showed that all data fields were filled in correctly, which indicates to automatic completion. It can be assumed that in case of manual entry, there would be some typing errors.

Name of the health care provider

Depending on the HCP, it is filled in by the HCP's information system or should be entered by the dentist. The data field was filled 100%, but in case of some HCPs, there were different variations of their names. Counting the HCPs who had sent dental records to the HIS based on the commercial register, the number was 324 while in counting by names of the HCPs, it was 340. This means that there are several variants of the name used in case of 16 HCPs, which may refer to manual entry or the difference in the way the names have been spelled in different systems. From the point of view of production of statistics, the occurrence of variations in name is not important because NIHD combines general information about HCPs from the list of health care providers managed by institute, based on the HCP's commercial registry code.

Commercial registry code of the health care provider

Depending on the HCP, it is filled in automatically by the HCP's information system or should be entered by the dentist. There were no errors here.

All fields are mandatory in the compiler data block of Health Information System's dental record document. There were no problems with filling of these fields that would prevent production of statistics.

2.2.3 Patient information

Patient's unique ID

Generated in an anonymised form by TEHIK upon data delivery. There were no issues with the data field.

Sex

Mandatory data field entered on the basis of a classification. The data field was completed 100% in all entries and corresponded to the classification.

Age

Data calculated by TEHIK based on the patient's date of birth, in years. Age values ranged from -91 to 159. The negative age value is probably due to incorrect entry of the patient's date of birth by the person filling in the document. Date of birth must be entered in the form DD.MM.YYYY.

There were 23 dental records in which the patient had a negative age, and the patient's age was absent in 1003 documents, accounting for less than one percent of all documents. These documents had been submitted to the HIS by 110 HCPs. On one document, the patient's age was 159, on the other, 115 and on 186 documents it was 105. The latter were all submitted by the same HCP. In these documents, there is also a doubt whether the patient's correct date of birth has been entered.

Real residence

Stating the patient's actual place of residence is optional. Detailed address of the patient's residence can be entered into HIS. In case of foreigners, the respective country code should be entered. At the moment, NIHD produces dental care statistics based on the place of business of the HCP. Data on actual residence would provide an opportunity to distinguish dental services provided to residents of Estonia from the services provided to foreigners and to produce dental statistics based on the patient's place of residence. For this analysis, patient's residence was required with the accuracy of municipality, but it appears from the data that TEHIK provided the data with the accuracy of a county.

The patient's residence was added in 187,704 documents, representing less than a quarter of all documents and visits. Different variations were used for the same places in case of people from foreign countries. For instance, Helsinki has been referred to as HELSINGI, HELSINKI or Helsinki. Two version were used for Estonian counties, in which one or two spaces were between the name of the country and the word "county", e.g., "Valga maakond" and "Valga maakond". From the point of view of statistics, such incorrect place of residence is not critical. It is possible to re-code such names during data analysis, but for NIHD, it would mean extra work.

The problem in the Health Information System's dental record' patient information block is incorrect entry of birth dates and the lack of data on birth and place of residence. Correct information about date and age are important to produce statistics. NIHD publishes data on children and adults separately, and it is also necessary to distinguish age groups to produce international statistics. If completion of the data field of patients' place of residence does not improve, the production of dental care statistics by patients' place of residence will not be possible or remains inadequate.

2.2.4 Dental care visit data

The dental care visit data at dental records includes the date of the visit, the diagnoses during treatment and information on the work performed. One dental record may include several dental care visits and several diagnoses and dental procedures made at different visits.

Visit date

From a statistical point of view, date of visit is a highly important data field. If this field is not filled in as required, the number of dentist visits cannot be counted. It is also important when the patient's age is calculated based on that date. In addition to national needs, dental care visit data are also provided for Eurostat/WHO/OECD joint questionnaire for international statistics. According to the instructions, the date of the visit should be entered in the format DD.MM.YYYY. In the HIS data analysed, the versions YYYYMMDD and YYYYMMDDHHMMSS had been used. The latter type for the date of the visit will not interfere with production of statistics. On some cases, four- and six-digit formats have also been used, probably by mistake.

The date of the visit was entered on 990,053 entries, i.e., on 26% of the cases (Figure 3), of which 944,121 entries the visit took place in 2016. In case of the ten entries, the year of the visit could not be determined. There were more than 45,000 entries in which the year of verification and receipt of the document was 2016 but the year of the visit was 2015 or earlier. Since the data extraction was done according to the year of the receipt of the HIS document (2016), the latter situation can not be considered an error since the dental record may also include patient's visits on previous years. There were 153 entries where the date of confirmation and receipt of the document was in 2016, but the year of the visit was in 2017. This is probably a mistake in entering the date of the visit. These documents mainly belonged to one HCP. If the Health Information System had logical checks on data fields, then such mistakes could be avoided.
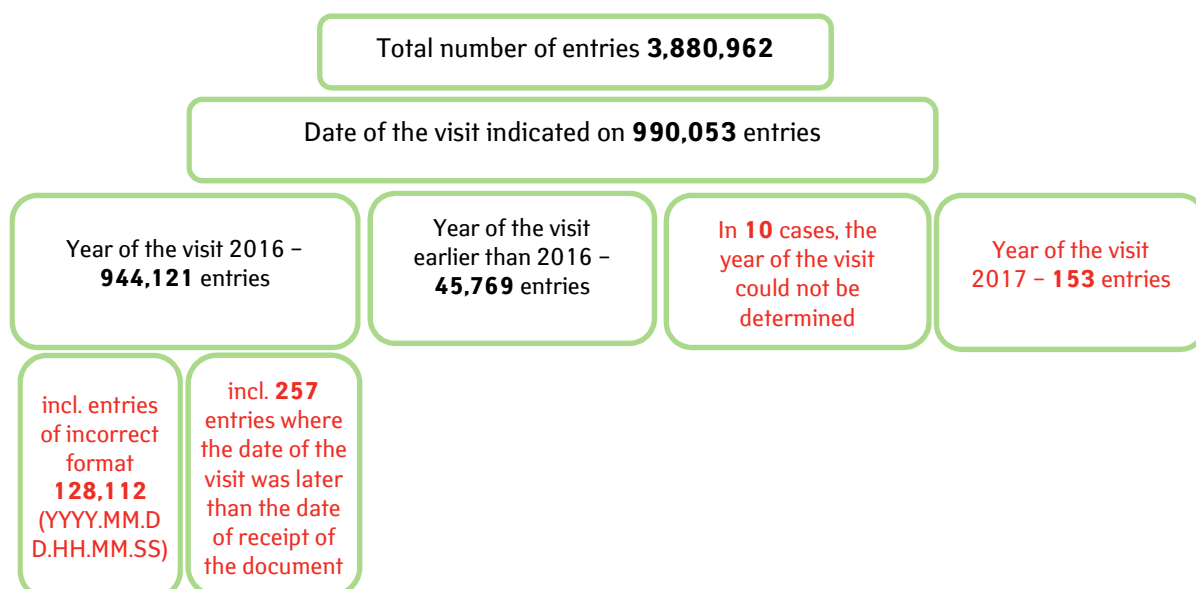
```
                  ┌──────────────────────────────────────────────┐
                  │   Total number of entries 3,880,962          │
                  └──────────────────────────────────────────────┘
              ┌──────────────────────────────────────────────────────┐
              │   Date of the visit indicated on 990,053 entries     │
              └──────────────────────────────────────────────────────┘
```

| Year of the visit 2016 – **944,121** entries | Year of the visit earlier than 2016 – **45,769** entries | In **10** cases, the year of the visit could not be determined | Year of the visit 2017 – **153** entries |

| incl. entries of incorrect format **128,112** (YYYY.MM.DD.HH.MM.SS) | incl. **257** entries where the date of the visit was later than the date of receipt of the document |

Figure 3. Distribution of the HIS dental record entries with the date of the visit, 2016

17 HCPs had entered no date of visit to any of the documents sent. These HCPs account for nearly **6%** of all HCPs who reported data to HIS. To the NIHD, these HCPs had submitted data on

more than 40,000 visits. The records also included entries where the date of the visit was later than the date of receipt of the document to the HIS. There were 257 such entries. Again, this error could be avoided if HIS or HCP's information system had logical checks on dates.

Number of visits, method A

Leaving out the entries where the year of the visit was impossible to identify and the duplicates sent, and counting only the entries where the year of the visit was 2016 (method A), there were almost a million visits based on the HIS dataset. In case of method A, duplicate documents were deemed to be the ones in which all the fields beside the document number were the same. There were 2,929 (0.3%) documents where the date of the visit was indicated and all the fields except the document number were the same.

Comparing the visits counted from the HIS by method A data and the number of visits reported to the NIHD, then visits sent to the HIS constituted 50% of the visits submitted to NIHD. When viewing the data coverage only across the institutions that had reported data both to HIS and the NIHD, the number of visits sent to HIS made up **64%**. Large differences appeared when comparing the number of visits by HCPs (Figure 4). Some of the HCPs had submitted significantly more and some significantly less visits to NIHD than the others. There were 41 HCPs whose number of visits fell into the 91–110% overlapping area. That means that **87%** of the HCPs had reported significantly more or less dental care visits to the NIHD.

The formula used to compare the data:

$$\textbf{Coverage}\ (\%) = \frac{Visits\ in\ NIHD\ reports}{Visits\ in\ HIS} \times 100$$

The overlap indicates how many times more visits are reported in the NIHD data compared to the data at HIS. If the overlap percentage was over 100, the number of visits in NIHD was higher (over-coverage), and if the overlap percentage was less than 100, the number of visits in NIHD was less than in the data in HIS indicates (under-coverage).
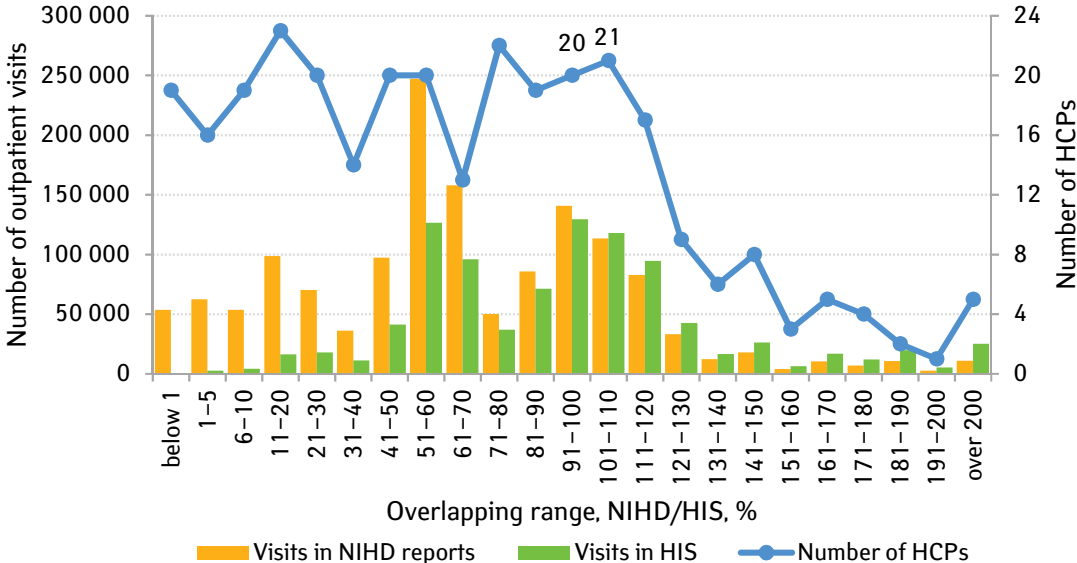

Figure 4. Overlapping of dental care visits and the number of institutions (%) in NIHD and HIS data sets by method A, 2016

Such large variations in the visit data raise the question of whether visits can be counted based on the method A described. By studying data from specific HCPs who have entered significantly more visits at HIS, it can be seen that the date of the registered visit does not always mean a single visit that actually took place, but the date is simply overbooked. Namely, HCPs mark the dates of visits very differently. In theory, it could be that the date of a visit is marked on a dental record on one visit once and the record is either sent or is resumed with the next visit. This is not done in actual practice. For example, there are 24 entries on a dental record in the HIS data set. There are 8 visits on that dental record, but **they all have the same date**. Each entry that indicates the date of the visit also has a compulsory field for the diagnosis of the main disease. According to these data, it appears that a single person had visited a dentist eight times in one day and received the same diagnosis eight times. It turns out that several teeth were repaired and the repair of each tooth was indicated in the dental records as a separate visit, i.e., the same visit date is marked for each tooth. In this case, the data field on the work performed contained a single filling for all these visits. Yet it is not realistic to believe that a total of eight teeth were repaired in one patient during a single visit. In these documents, it is not possible to determine when the visits took place: whether there were several visits within one day, eight teeth were repaired during a single visit, or the dental record was filled and sent to the HIS after completion of the treatment because the visit dates are all the same.

It is important to note that the date of the visit should be entered by the person filling the document, it does not appear automatically. The latter supports the assumption that in the example described above, dental records were formalised after the end of treatment and, as a result, all visits were entered with the same date.

In another example, a single document had 11 entries in the data set. Three of these entries contained the date of the visit, and there are two such dates. When counting the visits, the question arises as to whether three or two visits should be counted in this case. It is not possible to determine on the basis of the data whether two visits have taken place in one day, or if the date of the visit has been simply entered twice when sending the data to HIS.

The visit data overlapping varies widely by counties (Figure 5). Based on Method A, the lowest data overlaps are in Lääne-Viru, Hiiu and Põlva counties, where 31–41% of NIHD visits are also represented in HIS. Valga County has the best visit data overlapping – 100%. In interpreting the figure, it should be borne in mind that only the data of the institutions that have provided data to both HIS and NIHD has been taken into account when comparing the number of visits.
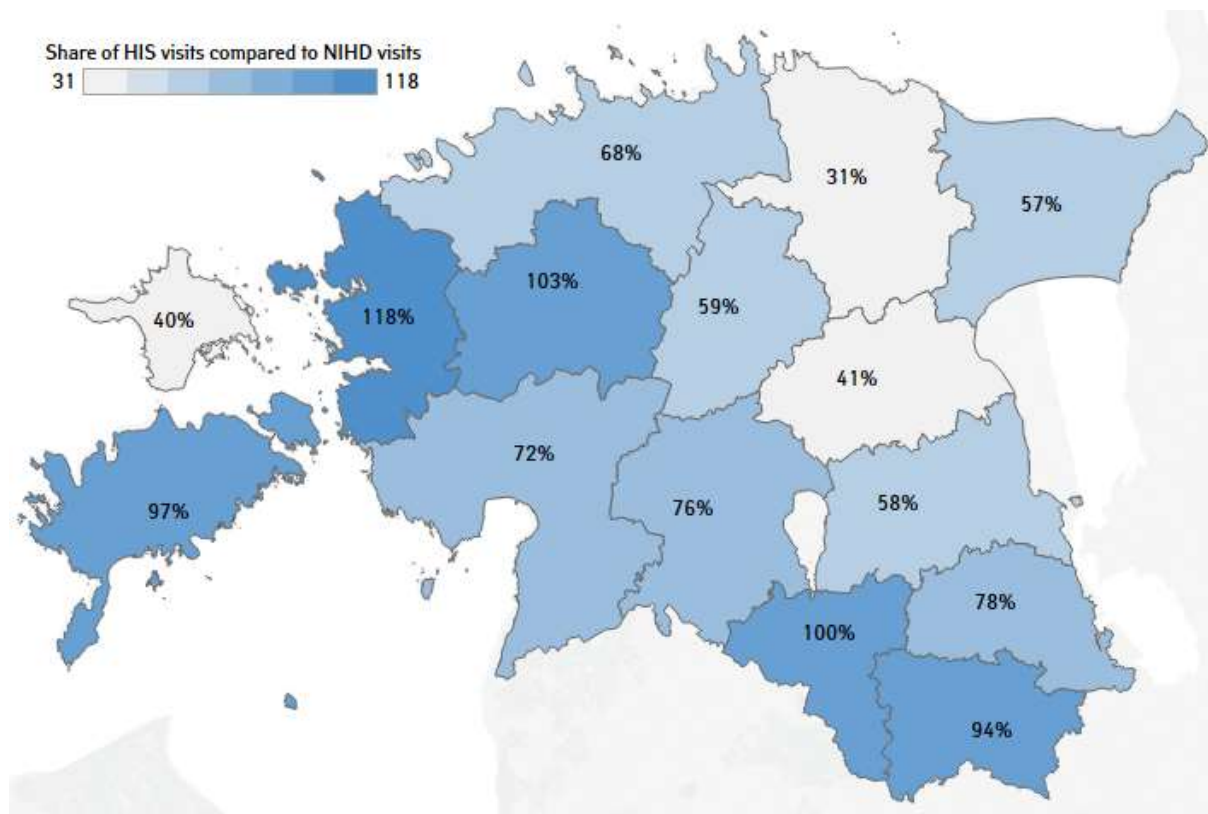
Figure 5. Share of the number of visits of the HCPs who submitted data to HIS in the number of visits reported to NIHD by county, method A, 2016

Number of visits, method B

Using method B for calculating the visits, leaving out the dental formula upon determining duplicate documents, the data overlapping of the number of visits decreased by 10% (over 187,000 visits less) compared to method A. The visits sent to HIS represent 40% of the visits reported to NIHD in case of method B. When viewing the data coverage only on the institutions that had reported data both to HIS and the NIHD, the number of visits sent to HIS made up **51%**.

Comparing the drawing on the basis of method A (Figure 4) and method B (Figure 6), it can be observed that in case of method B, the last part of the graph reduces to zero. This indicates that it is possible that the date of the visit is marked with each tooth fixed and diagnosed and/or treatment procedure, or a dental record is submitted at the end of the treatment reporting all the visits with the same date. In case of method B, the number of visits dropped for 73% of the HCPs; the number remained the same for the rest of the HCPs.
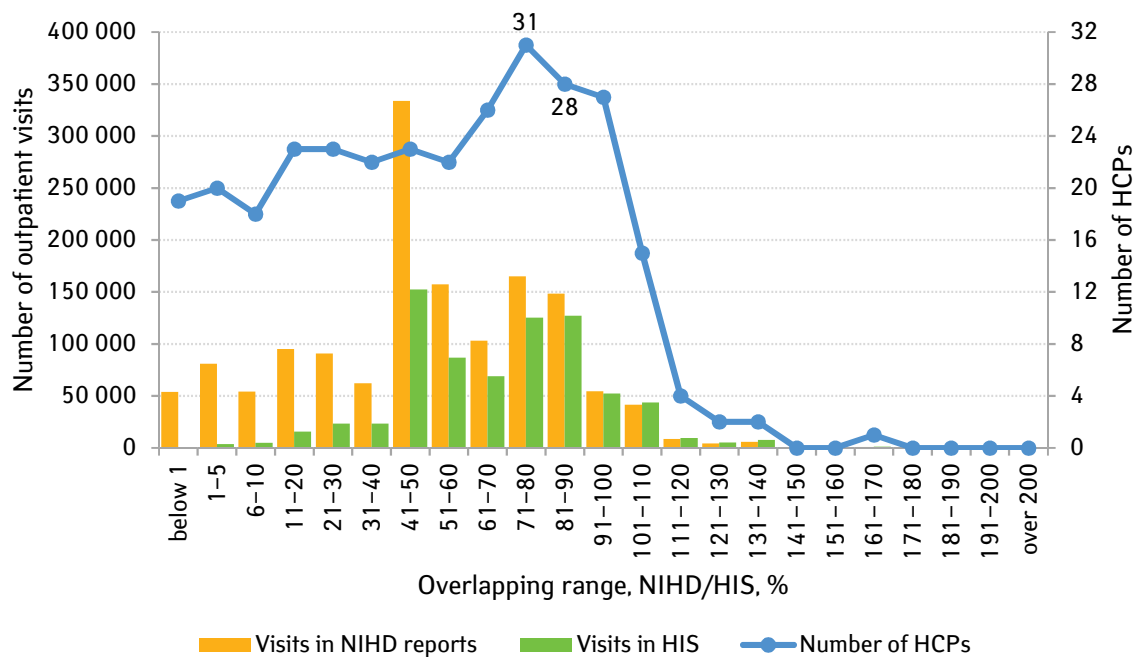
Figure 6. Overlapping of dental care visits and the number of institutions (%) in NIHD and HIS data sets by method B, 2016

In the case of HCPs that had much more visits in the HIS database in case of method A than that of visits in the NIHD data set, then in case of method B had a decrease in visit overlapping by 24–285 percentage points. Only one institution maintained the same number of visits.

Data overlapping is very different by counties also when visits are counted based on method B (Figure 7). Similarly to method A, the lowest data overlaps are in Lääne-Viru, Hiiu and Põlva counties, where 31–36% of NIHD visits are also represented in HIS. Rapla County has the best visit data overlapping – 89%. At Valga County that had a 100% overlapping when calculated by method A, the same indicator drops by 19 percentage points in case of method B. Looking at the indicators of Lääne County, where in case of method A, 18% more visits were submitted to HIS than to NIHD, the number of visits in case of method B has dropped considerably – the number of HIS visits makes up 79% of the visits reported to NIHD. This indicates that at Lääne County, reporting each repaired and diagnosed tooth as a visit is a common practice.
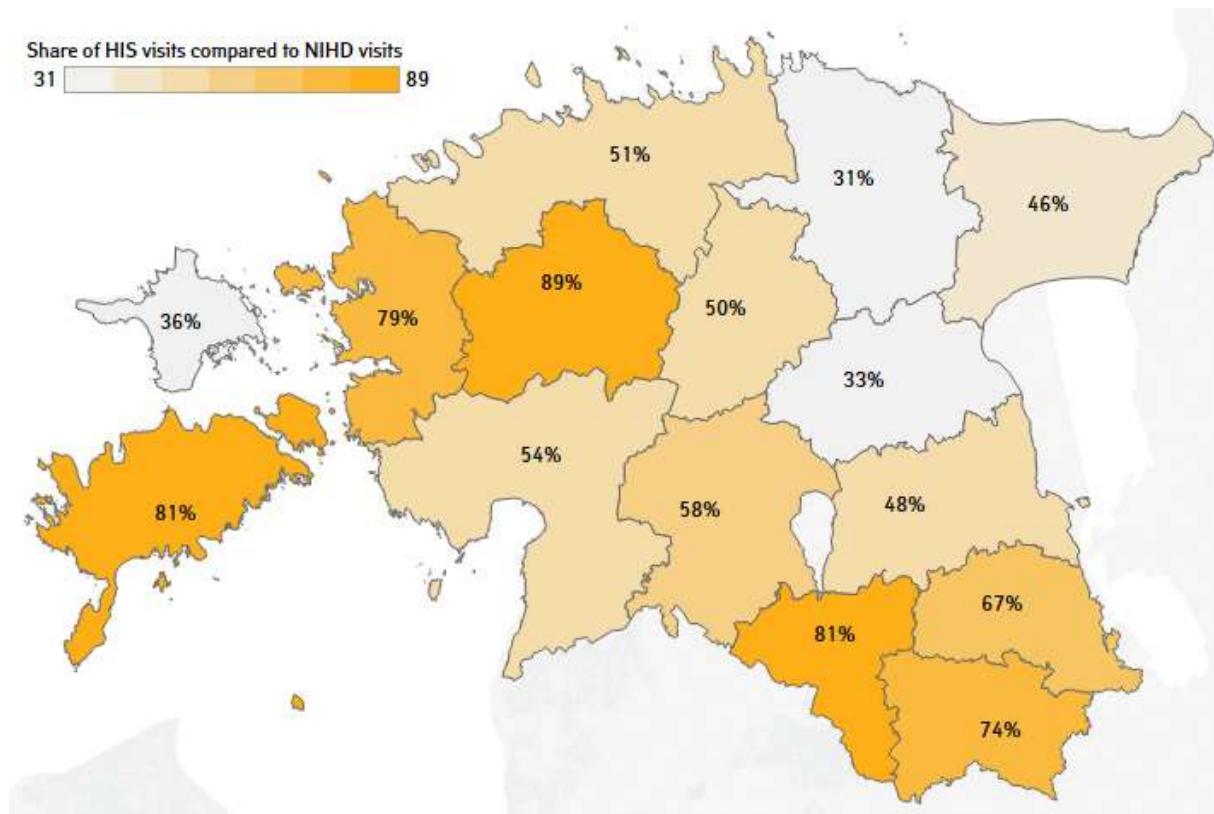
Figure 7. Share of the number of visits of the HCPs who submitted data to HIS in the number of visits reported to NIHD by county, method B, 2016

In case of method B, all the dental records in which the values of all fields, except for the document number and the principal diagnosis dental formula, are the same, are considered a duplicate. This rules out double counting of the visits where, for example, tartar is removed from eight teeth and the date of the visit added for each one of them, causing overestimation of the number of visits in statistics. The disadvantage of method B is that if there really were two visits per day, during which two different teeth were fixed, it is still counted as one visit. In case of method A, it was considered two visits because the dental formula of the principal diagnosis was different and did not count as a duplicate. There is also an issue with documents that were prepared and sent to the HIS after the end of treatment. For instance, the case in which the dental records showed 8 teeth fixed in a single visit. In this case, the dates of confirmation, receipt and visit of the document, as well as all other data fields were the same, only the dental formula was different. However, since the variable of the dental formula was left out when considering duplicates, only one visit was included in the statistics.

Counting dental care visits based on the dental records of the HIS is currently very complicated, if not impossible. To count in the visits, an extra field needs to be added to dental records on the overall number of registered visits.

All data fields in the main disease block are mandatory.

Code and name of the diagnosis according to ICD-10

The code and name of the diagnosis of the main disease is entered to dental records by a health care professional and in accordance with ICD-10.

Diagnosis code was marked with one to five digits ICD-10 codes. From the point of view of producing statistics, using only one or two-digit codes is insufficient. For instance, codes K and K0 have been entered. In statistics, it is necessary to distinguish between the number of periodontal care visits and the number of teeth with root canal treatment, calculated according to the three-digit ICD-10 code K05 (gingivitis and periodontal diseases) and K04 (diseases of pulp and periapical tissues).

The three-digit code K05 was marked as the principal diagnosis in more than 73,000 and the code K04 in almost 170,000 entries. The diagnosis of the main disease was a two- or one-digit code K0 or K for a total of 1,148 entries, and it was not possible to distinguish between a periodontal care visit, root canal treatment, or some other K chapter (digestive diseases) diagnoses.

There were also codes not included in the current official version of ICD-10 as the principal diagnosis, for example G47.63 (named "bruxism"), K03.89 (called "hypersensitive dentin"), or codes that failed to match the name of the diagnosis to a code such as Z01.9 (marked as "did not appear for visit"). The correct definition for Z01.9 according to ICD-10 is "special examination, unspecified".

There have also been principal diagnoses that are difficult to link with dental care. For instance, the principal diagnosis has been K41 ("femoral hernia"), G47.3 ("sleep apnoea"), Z01.4 ("gynaecological examination") or Z01.3 ("examination of blood pressure"). In the last example, there has probably been an error in entering the code as it is preceded by code Z01.2 in ICD-10, marking dental examination which would be the correct code to use if the patient had no other disease.

All of these codes were used on only a few occasions, but as such data is still in the data set, they together already form an important part and can affect the overall statistical results.

In all cases, the name of the principal diagnosis was marked in some form when the diagnosis code was present. Both Estonian and Latin names were used. Theoretically, the incorrectly entered codes could be corrected based on the name of the principal diagnosis, but it cannot always be done in practice. Namely, in some cases, the name of the diagnosis also included the one-letter name. In addition, work with text fields is extremely time consuming.

Verbal clinical diagnosis

The verbal clinical diagnosis of the main disease is entered by the health care professional. These data fields contained similar errors than the names of the principal diagnoses. From the statistics point of view, this field is not important, but it plays an important supporting role if the inserted main disease code is faulty.

Statistical type of diagnosis

The statistical type of diagnosis of the main disease is entered as a code by the health care professional in accordance with the current classification: new case (+), recurring case (-), initial diagnosis (0).

The statistical type of diagnosis is present when there is a principal diagnosis. Values +, −, 0 and "new case" had been entered. In 99% of the cases, the statistical type of the principal diagnosis was + or "new case".

This data field is not important in the context of the dental care statistics currently produced by the NIHD, since it has not been collected for the past 20 years in the report. In the future, we need to distinguish between new and recurring cases. In such case, there may be a problem because it is clear that the data field is not being filled correctly.

The principal diagnosis is predominantly a new case, which should mean a diagnosis that is made for the first time in the patient's life. It is unlikely that 99% of the principal diagnoses are first cases. Rather, it indicates that the data field is only filled in the same way because it is mandatory or the data field is set to "+" by default. For instance, the principal diagnosis is code Z01.2, denoting dental examination, and the statistical type of the diagnosis is "new case". This may be true for children of up to few years old, but this diagnosis has also been used for patients 2–86 years old. Older people are unlikely to have been visiting a dentist for the first time. Code K00.0 marking anodontia, for example, has also been reported as a new case for the teeth of the same person. In this case, the statistical type of the diagnosis is clearly incorrect.

Dental formula

Based on the classification, the number of the tooth on which the work was done, is entered. According to the classification, the values entered should be either two-digit numbers (e.g. 11) or a combination of L and two numbers (e.g. L61).

The HCPs had filled in the data field in very different ways. Values provided in the classification, as well as various symbols (!, #, $; %,), *, ?, /, @ etc.), single letters (A, I, J, R, T etc.) and single numbers had been added. This suggests a great confusion in filling the data field. At the same time, the entries where the dental formula was filled in and the values not present in the classification were used only made up 0.04% of the data set.

This field, like all other data fields in the main disease block, should be mandatory according to the instructions. At the same time, the value of the dental formula was missing in 191,227 entries, which make up one fifth of the entries that contained the principal diagnosis. When a classification was used, one to several dental formulas were indicated on a **single data field**.

This is an important data field. If several teeth were repaired during one visit, it should be possible to link the diagnosis to a specific tooth based on the dental formula. If the dental formula is not stated, this information will be incomplete. For example, it is not possible in this case to count the teeth that have received root treatment, on which NIHD collects national statistics.

In case of the block on the main disease diagnosis, the problems related to production of statistics include indication of diagnosis codes with insufficient level of detail, misinterpretation of the statistical type of the diagnosis, and stating non-characteristic values for the classification in case of dental formula.

In the case of a complication of a main disease diagnosed, it is mandatory to fill the subsection on complications from the main disease in the dental records.

Code and name of the diagnosis according to ICD-10

The code and name of the diagnosis of the complication from the main disease is entered by a health care professional and in accordance with ICD-10.

Diagnosis of a complication from the main disease had been registered by five HCPs in a total of 112 entries. In case of these entries, the four-digit code was correctly entered, corresponding to the classification and being related to dental care. In all cases, the name of the diagnosis of the complication was added in some form if the complication diagnosis code was available.

Since only five HCPs had provided the data, the HIS data does not allow actual statistics on complications.

Verbal clinical diagnosis

The verbal clinical diagnosis of the main disease is entered to the free text field by the health care professional. The verbal clinical diagnosis of the complication of a main disease corresponded to the name of the complication diagnosis code in all cases.

Dental formula

As the dental formula for complication of the main disease, the number of the tooth to which the corresponding work was performed is entered according to the classification. According to the classification, the values entered could be either two-digit numbers (e.g. 11) or a combination of L and two numbers (e.g. L61).

In all cases where the data field was filled, it complied with the classification and a single dental formula was entered per entry. According to the instructions, it is a mandatory data field, but in one-third of the cases regarding diagnosis of complications from the main disease, the dental formula of the complication from the main disease had been left out.

There were no obstructions to the production of statistics in the block on complications from the main disease. The issue only concerns the lack of data on complications from the main disease.

In the case of a concomitant disease diagnosed, it is mandatory to fill the subsection on the concomitant disease in the dental records.

Code and name of the diagnosis according to ICD-10

The code and name of the diagnosis of the concomitant disease is entered by a health care professional and in accordance with ICD-10. One, four and five-digit ICD-10 codes were used in the data set. One fifth of the HCPs supplying data to HIS provided data on the concomitant disease.

If the code of the diagnosis was marked, the name according to the ICD-10 was also added in all cases. Both Estonian and Latin names were used.

Verbal clinical diagnosis

The verbal clinical diagnosis of a concomitant disease is entered to the free text field by the health care professional. The data field was filled in all cases when there was a code for the diagnosis of the concomitant disease.

Statistical type of diagnosis

The statistical type of diagnosis of the concomitant disease is entered by the health care professional in accordance with the current classification: new case (+), recurring case (-), initial diagnosis (0).

Values +, - and 0 had been entered. The statistical type of diagnosis of a concomitant disease was not stated in 98% of the entries in which the diagnosis of the concomitant disease had been marked. Leaving out the statistical type of diagnosis of a concomitant disease raises questions since this field should be mandatory.

Dental formula

As the dental formula for a concomitant disease, the number of the tooth to which the corresponding work was performed is entered according to the classification. According to the classification, the values entered could be either two-digit numbers (e.g. 11) or a combination of L and two numbers (e.g. L61).

If the data field was filled in, it corresponded to the classification in all cases. Although a four-digit dental formula was entered in some of the data fields, it was probably just a case of accidentally leaving out a comma between two tooth numbers. According to the instructions, this is a mandatory data field, but in 13% of the cases with concomitant disease, this dental formula was left out.

From the block of concomitant diseases, production of statistics could be hindered by not stating the statistical type of the diagnosis of the concomitant disease and the dental formula. Both fields are mandatory, but how could the documents be submitted at HIS when this part was left empty.

Data on works performed

Data on works performed shows the operations carried out with the respective teeth during the visit.

Code and name of the work performed

The code and name fields of the work performed require the code of the performed work to be entered according to the EHIF's list of health care services or the corresponding alternative classification. Allowing the use of different classifications for a single identifier refers to a non-standardized variable that should not be allowed if a unified national information system is to be created.

The name of the work performed was entered in almost all cases when there was a code for the work performed. The name was missing in just three entries, although the code of the work had been entered.

NIHD has no overview of the alternative classifications The list of health care services of the Estonian Health Insurance Fund does not cover the entire list of services offered by HCPs and so the HCPs have coded the missing services themselves. Data analysis shows that very different encodings have been used. Examples of the used codes can be seen in Table 1.

Table 1. Examples of the used codes for works performed in HIS dental records, 2016

Codes used				
00	5437a	kl	p01	prot.par
000.1	5440–1	kl.t	paro1207	pu101
1	5467.9.	kl2	pk301	r.t
020803-A	8,1	m904	pr	r1
10-a	ALE2	mkr	pr.5	rvt
5435-AA	AN	oo201	pr.par.	s.a
5435A	hp45	opr	pr1	O-1a

There are a total of 5,616 codes on the works performed used in the data set. From the current list of EHIF health care services, an equivalent can be found for 328, which represents 6% of all codes used. Looking at the names of the codes that had a match in the EHIFs list of health care services, there seem to be some service codes that have a weak or incomplete relationship with dental care (Table 2). At the same time, the name given by HCPs gives the code a meaning in dental work.

Table 2 shows that HCPs are using classifications present in the EHIF's list of health care services, but these have an entirely different meaning. There are also codes that mark different works or services in case of different HCPs. This makes production of statistics very difficult when the identification is not standardised. Using a number of different methods for processing text fields, it is theoretically possible to distinguish between the contents of a work if the same code means different operations in different institutions, but can be done as a part of a one-off study rather than for regular national statistics. Production of statistics on the basis of text fields is very complicated and time consuming. Continuing with collecting such mixed-type data and spending resources is also not practical.

Table 2. Examples of used codes and names of the work performed in HIS dental records, 2016

EHIF code	EHIF name	Name in data set
3101	Deciding whether to place an organ transplant on the waiting list or review the decision by a multidisciplinary expert committee	Light-curable filling of one surface (medium)
4001	Whole blood (1 dose 468–558 ml)	Use of cofferdam kit; lightening gel, removal of single-root tooth or a root
4013	Blood plasma, child's dose (1 dose up to 120 ml)	Dental fiber splint (2 teeth)
4021	1 BC thrombocyte concentrate 45–85×109	denture instalment 100
4022	Reconstituted whole blood for exchange transfusion	denture instalment 10
7056	Physiotherapy in a pool, individual (duration 30 min)	Laser reconstruction of gum line on one tooth
7057	Physiotherapy in a pool, group (duration 30 min)	Herpes or stomatitis aphthosa treatment, 1 session
7059	6-minute walking test	Removal of a fibroma or hyperplasia

This is more complicated when several different meanings are assigned to the same code within the HCP. This is especially prominent in 34 HCPs who had used the code 0000 in a total of almost 140,000 times and given it more than 1,300 different meanings. Some examples of the different names for code 0000 can be seen in Table 3.

Table 3. Examples of different names for code 0000 in dental records, 2016

Name of the work performed	
Temporary plastic crown from a laboratory	Cavity preparation
Temporary cement	Consultation, compiling a treatment plan, stating a status
Basis under a filling	MetaFix
First visit, compilation of status, use of intracam	ProTemp temporary crown
Fiber post	Soda cleaning of upper and lower jaw
Wound or mucous membrane cleaning, cauterisation with medicine	Zirkonium crown
Tooth or root removal	Teeth Whitening Strips
Dentist visit	Cementation with light-curable filling
Removal of tartar	Removal of filling
Removal of tartar in one region	Filing polishing
Dental examination	Whitening Perfict Bleach Office+
Interdental brush	Light-curable composite filling
Composite sheening of two teeth	Visit 15 min

Looking at the data field values for the works performed, it looks like a list of activities and details necessary for paying for dental care (Table 4). All of these values have been added to some extent, but the list is not complete and all of these together constitute a significant part of the data set in the end.

Table 4. Examples of names of works performed on payment details in HIS dental records, 2016

Name of the work performed		
rush work %	visit	paid from prepayment
denture instalment 100	first visit fee	prepayment without VAT
denture instalment 10	toothbrush for dentures	additional laboratory costs
1212 Visit fee	cost of service	laboratory services

Dental formula

As the dental formula for works performed, the number of the tooth on which the corresponding work was performed is entered according to the classification. According to the classification, the values entered could be either two-digit numbers (e.g. 11) or a combination of L and two numbers (e.g. L61).

The data field had been left empty throughout the data set. Unfortunately, it is not clear whether it is empty because none of the HCPs had filled this field, the data field cannot be filled or there is an error in the information system. Currently, the tooth and the work done on it cannot be linked based on the data at HIS.

The biggest obstacle to the production of statistics in the block on the works performed is the use of very different classifications and the complete non-fulfilment of dental formula data of the work performed. At the moment, only 6% of the codes used can be found in the EHIF list of health care services and some of them are not related to dental care, i.e. the list used is not the one of EHIF, but HCPs have used the codes in listed in there in a different meaning in their own alternative classifications.

2.2.5 Data on clinical examination in dental care

Cariologic tooth indexes

Cariologic tooth indexes are the DMFT and dmft index. Both data fields were left empty throughout the data set.

These are important data fields. The average DMFT value of 12-year-olds is required by OECD in its international data inquiry. Estonia has published information on its DMFT index twice, the last time was in 1998. Eurostat has also expressed interested in when Estonia would have an updated DMFT index data. Eurostat has been told that the DMFT index data should be based on the Estonian Health Information System, but it is yet uncertain when the data could be used.

Code and name of the periodontological diagnosis

The code and name of the diagnosis of the periodontological diagnosis is entered by a health care professional and in accordance with ICD-10. The block on periodontological diagnoses had been filled in by a single institution with three entries that were all correct.

Gingivitis data

The status of gingivitis is assessed based on the valid classification. On gingivitis, the disease diagnosis code and name are entered in accordance with ICD-10, and a verbal clinical diagnosis and dental formula indicating which tooth was evaluated.

Only one HCP had submitted the data on gingivitis in dental records, having made 41 entries. This is the same HCP who had filled the section on periodontological status. The current classification was used to determine the extent of gingivitis: absent, marginal, local, generalised. The ICD-10 code stated was K05.1 or K05.2 and the name of the clinical diagnosis corresponded with the code of the diagnosis.

Data field on the degree of gingivitis was filled incorrectly – classification on the extent of gingivitis was used. According to the dental records instructions, this should be stated with a number – 0, 1, 2 or 3. The gingivitis dental formula was entered according to the current classification.

The problem in the clinical examination block was the lack of data – it is not enough to produce statistics. Only one HCP submitted the data, that on the minimal extent.

2.2.6 Data on summary of patient's treatment

Reason for ending treatment

When necessary, the reason for discontinuation of treatment is entered based on a valid classification: completed treatment (1), consultation only (2), suspended due to complications (3), suspended at the patient's request (4), and interrupted (5).

The data field was filled by ten HCPs with a total of 1,638 entries. Valid classification had been used. From the point of view of production of statistics, filling this data field is not critical because it is not currently used. However, these data fields may prove necessary in the future in case of using detailed data.

Performance of sanitation

Sanitation is indicated by either "yes" or "no" according to whether or not the oral cavity was sanitised or treated. 1,572 entries had been made on the data field, of which 92% had the value TRUE, meaning that the oral cavity had been treated. A total of eight HCPs had filled in the field. These are the same HCPs who had filled the data field on the reasons for ending treatment.

As with the previous clinical examination data block, the lack of data is a problem also with patient treatment summaries. 2% of the HCPs who had submitted data on HIS had filled in the field.

3 Conclusions

❖ 3/5 of dental care providers send dental care data to HIS.
❖ This is a fraction of the actual dental work volume.
❖ Some variables are not standardized in HIS, which makes the use of a unified national information system unsuccessful.
❖ HIS often receives substandard data.
❖ HIS does not provide the additional variables needed to produce national statistics. These variables are either completely missing or partial.

4 Recommendations

Currently, the country only monitors the submission of HCP data to HIS, no matter how many documents and variables are entered to the system. Data quality, which is very important to produce statistics, has also been neglected. Our wish is that HIS would start reflecting the real situation in the Estonian health care system, the data would support patients' treatment and would be usable for both regular statistics and other studies.

For the above, the following steps are necessary:

❖ By the **Estonian Dental Association**, to encourage and motivate HCPs providing dental care services to submit data on HIS and fill in the dental records (not just submit data on outpatient epicrisis) in time and in a correct manner.
❖ By **TEHIK**,
  ▪ to standardise the variables and classifications. Also take responsibility for data quality. Mandatory fields need to be mandatory, that is, filled correctly;
  ▪ carry out regular trainings for HCPs on completing dental records and establish customer service to respond to real-time questions occurring during the reporting process;
  ▪ start or intensify cooperation with software developers to clarify the content of the variables and verify achievement of the goal.
❖ By **EHIF**, include a criterion on submission of high-quality data to HIS to the service prioritisation model of the contract for provision of services, since it is part of the provision of services in the country.
❖ By the **Health Board**, carry out constant monitoring on the provision of high-quality data to HIS.
❖ By **NIHD**, to continue with data quality studies, bringing out the strengths and weaknesses of the system. For this, it is important to have regular dental care data extracts, similarly to inpatient and outpatient care data, to regularly monitor the quality of data submitted to HIS.
❖ Cooperation of **NIHD** with the **MoSA, TEHIK** and **Estonian Dental Association** to develop the correct methodology for calculating the number of dentist visits or to add a data field to the dental records for a summary of the total number of visits, as it is solved in outpatient epicrisis.
  Since NIHD has no right to combine data from different data sources on a regular basis, using a personal identification code to enrich and enhance the quality of the data, **TEHIK** needs to develop a system of pseudonymisation to enable linking the values of different variables from various sources.

Health and health care statistics:

- **Health statistics and health research database**
  http://www.tai.ee/tstua

- **Website of Health Statistics Department of National Institute for Health Development**
  http://www.tai.ee/en/r-and-d/health-statistics/activities

- **Dataquery to National Institute for Health Development**
  tai@tai.ee

- **Database of Statistics Estonia**
  http://www.stat.ee/en

- **Statistics of European Union**
  http://ec.europa.eu/eurostat

- **European health for all database (HFA-DB)**
  http://data.euro.who.int/hfadb/

- **OECD's statistical databases (OECD.Stat)**
  http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT