Modelling declarative knowledge from a clinical guideline on gestational diabetes (GDM) with temporal abstraction knowledge concepts (TAK):

A graph theoretical & statistical analysis of manually & automatically generated models

Final Report

Bettina Wimmer and Katharina Kaiser
University of Applied Sciences St. Pölten
bettina.wimmer.mail@gmail.com
katharina.kaiser@fhstp.ac.at

This project was funded by the Austrian Science Fund (FWF) as part of the project "Support for the development & transformation of clinical guidelines and procedures" TRP71-N23 (2010-2015).

Introduction

The aim of this project was to model declarative knowledge in a clinical guideline on gestational diabetes with adequate TAK (Temporal Abstraction Knowledge) concepts. TAK serves to model temporal relations between human properties, medical events and findings. After the modelling a comparison between an automatically created model and the manual model was accomplished. Statistical and graph theoretical approaches revealed differences in the structure of the two models and the concepts they used.

Modelling declarative knowledge of a clinical guideline on GDM

> *"Gestational diabetes [GDM] is defined as carbohydrate intolerance that begins or is first recognized during pregnancy."*
>
> -- Gestational Diabetes Guideline (CSPT)

Method

The starting point was a clinical guideline on gestational diabetes which advises medical personnel of how to diagnose and support women affected by gestational diabetes (GDM). All declarative knowledge in this guideline had previously been identified and marked as "Gold standard" by Katharina Kaiser. In contrast to procedural knowledge (knowing how to perform a specific activity) declarative knowledge is defined as factual knowledge and information that someone knows. The guideline with the marked declarative knowledge can be found in the appendix.

The first task of this project was to manually model the declarative knowledge with TAK, a markup language schema based on XML. For each marked declarative knowledge string the

appropriate TAK concept types were chosen and logically combined to form a model of this phrase. All these models were saved in one XML file.

The second task was to analyze and compare these models to models generated by the computer program LASSIE. These models are based on the same clinical guideline. However, they represent declarative knowledge identified by LASSIE.

Thus, the two models are based on the same guideline though not necessarily on the same declarative knowledge strings.

Finally, the manually created models were compared to LASSIE's models generated from declarative knowledge marked by a modeler.

Explanation of the concepts and their usage: TAK/XML

TAK (temporal abstraction knowledge) is an XML schema to model temporal relations particularly in a medical context. Five different concepts were used for our models. For each concept we later searched the UMLS Metathesaurus data base1 for an appropriate term. This helps to compare and classify a concept according to standardized medical terms. For example, the concept "woman" is a population group concept with the CUI number C0043210.

Primitive concepts

Raw nominal, numeric or ordinal concepts were used to model basic medical events (e.g. results of a specific measurement) or personal characteristics (e.g. sex, former pregnancies). An example for a numeric concept would be the result of a measurement (e.g. fasting blood glucose concentration). Nominal concepts are particularly useful to model Boolean concepts. This could be a medical event (e.g. prediabetes) or a specific finding (e.g. macrosomia) which is either true or false. Ordinal concepts are used for findings or measurements on an ordinal scale (e.g. high/low). They are rather rare, because most of the measurements have a numeric output and can thus be used as a numeric concept. Figure 1 shows an example for a nominal concept.

```
<RawNominalConcept id="100001" name="sex" description="sex" type="RawNominal">
    <LocalPersistence goodBefore="0" goodAfter="0" timeUnit="Seconds"/>
    <TemporalSemanticProperties concenatable="false" downwardHereditary="false"
        upwardHereditary="false" forward="false" backward="false" solid="false" gestalt="false"/>
    <Synonyms/>
    <StandardTerms>
        <StandardTerm vocabulary="UMLS" key="sex" CUI="C1522384"/>
    </StandardTerms>
    <NominalAllowedValues>
        <NominalStringValue value="female"/>
        <NominalStringValue value="male"/>
    </NominalAllowedValues>
</RawNominalConcept>
```

Figure 1. Example: "Sex" is modelled as a raw nominal concept with values female/male according to the sex of the patient. There was a UMLS entry for the key word "sex" with the CUI number C1522384 (cf. section on CUI concepts).

Value abstraction concepts

These concepts are always abstracted from a predecessor concept. This means that they are mostly derived from the value of primitive concepts. In this example the concept “woman” is derived from the concept “sex” with the value “female”. Note: This is one of the rare cases where there is a UMLS entry for a value abstraction concept. Usually value abstraction concepts model rather complex phrases (e.g. “women with high risk for GDM who are unable to take an oral glucose load”) and thus do not represent a CUI concept.

```
<ValueAbstractionConcept id="100002" name="woman" description="woman" type="ValueAbstraction">
    <LocalPersistence goodBefore="0" goodAfter="0" timeUnit="Seconds"/>
    <TemporalSemanticProperties concenatable="false" downwardHereditary="false"
        upwardHereditary="false" forward="false" backward="false" solid="false" gestalt="false"/>
    <StandardTerms>
        <StandardTerm vocabulary="UMLS" key="woman" CUI="C0043210"/>
    </StandardTerms>
    <NecessaryContext/>
    <DerivedFrom>
        <ConceptID_Name ConceptID="100001" ConceptName="sex"/>
    </DerivedFrom>
    <DerivedFromComponents>
        <DerivedFromComponent>
            <ConceptID_Name ConceptID="100001" ConceptName="sex"/>
            <componentValueConstraints value="female" Operator="Equals"/>
        </DerivedFromComponent>
    </DerivedFromComponents>
</ValueAbstractionConcept>
```

Figure 2. Example of a ValueAbstraction concept.

Pattern concept

Pattern concepts are used to model recurring events or temporal intervals. They always abstract from one or more other concepts. We give an example for the pattern concept “ketonuria negative for two weeks” which is derived from the value abstraction concept

```
<PatternConcept patternComponentsRelation="AND" id="100017"
    name="ketonuria negative for two weeks" description="ketonuria negative for two weeks" type="Pattern">
    <LocalPersistence goodBefore="0" goodAfter="0" timeUnit="Seconds"/>
    <TemporalSemanticProperties concenatable="false" downwardHereditary="false"
        upwardHereditary="false" forward="false" backward="false" solid="false" gestalt="false"/>
    <Synonyms/>
    <StandardTerms/>
    <NecessaryContext/>
    <DerivedFrom>
        <ConceptID_Name ConceptID="100016" ConceptName="ketonuria negative"/>
    </DerivedFrom>
    <PatternComponents>
        <PatternComponent>
            <ConceptID_Name ConceptID="100016" ConceptName="ketonuria negative"/>
            <componentLocalConstraints>
                <componentValueConstraints value="true" Operator="Equals"/>
                <componentTimeConstraints referencePositionNumber="last"
                    referenceBoundaryPoint="End" timeUnit="weeks" MinDuration="2"
                    MaxDuration="2">
                    <ReferencePoint ConceptID="0"/>
                </componentTimeConstraints>
            </componentLocalConstraints>
        </PatternComponent>
    </PatternComponents>
    <PatternTemporalPairwiseConstraints/>
    <PatternPeriodicConstraints timeUnit="Seconds" minCardinalityValue="0"
        maxCardinalityValue="0" cardinalityValue="0"/>
</PatternConcept>
```

Figure 3. Example of a PatternConcept.

“ketonuria negative” with a temporal constraint of “two weeks”.

Context concept:

Context concepts (see Figure 4 for an example) are derived from a second concept and generated for a specific time frame provided from a third concept. However, this complex setting is rather rare. The example shows the context concept “pregnant woman” which is derived from the concept “woman” and generated while the concept “pregnancy” is true.

```
<ContextConcept id="100005" name="pregnant woman" description="pregnant woman" type="Context">
    <LocalPersistence goodBefore="0" goodAfter="0" timeUnit="Seconds"/>
    <TemporalSemanticProperties concenatable="false" downwardHereditary="false"
        upwardHereditary="false" forward="false" backward="false" solid="false" gestalt="false"/>
    <Synonyms/>
    <StandardTerms/>
    <DerivedFrom>
        <ConceptID_Name ConceptID="100002" ConceptName="woman"/>
    </DerivedFrom>
    <GeneratedFrom timeUnit="Seconds" timeGap="0" boundaryPoint="Start">
        <ConceptID_Name ConceptID="100004" ConceptName="pregnancy"/>
        <componentValueConstraints value="true" Operator="Equals"/>
    </GeneratedFrom>
    <GeneratedUntil timeUnit="Seconds" timeGap="0" boundaryPoint="End">
        <ConceptID_Name ConceptID="100004" ConceptName="pregnancy"/>
        <componentValueConstraints value="true" Operator="Equals"/>
    </GeneratedUntil>
</ContextConcept>
```

Figure 4. Example of a Context concept.

Logic Abstraction concept:

The logic abstraction concept (see Figure 5 for an example) is used to model logic contexts such as AND or OR between two or more concepts. Thus it is derived from at least two other concepts. In this example the concept “Woman with GDM” is derived from the concepts “woman” AND “GDM”. The concepts are identified by their unique ID.

```
<LogicAbstractionConcept ComponentsRelation="AND" K="0" id="100000"
    name="Woman with GDM [Logic]" description="Woman with GDM [Logic]" type="LogicAbstraction">
    <LocalPersistence goodBefore="0" goodAfter="0" timeUnit="Seconds"/>
    <TemporalSemanticProperties concenatable="false" downwardHereditary="false"
        upwardHereditary="false" forward="false" backward="false" solid="false" gestalt="false"/>
    <Synonyms/>
    <StandardTerms/>
    <NecessaryContext/>
    <DerivedFrom>
        <ConceptID_Name ConceptID="100002"/>
        <ConceptID_Name ConceptID="99992"/>
    </DerivedFrom>
</LogicAbstractionConcept>
```

Figure 5. Example of a LogicAbstraction concept.

Categorization of declarative knowledge:

The idea of this section is to sort the declarative knowledge strings of a medical guideline (such as the guideline on GDM) into classes, which can be modelled in a similar way. The categorization enables a collection of possible models for a specific category. We established nine different categories. Evidently, this list is not conclusive and serves rather as a starting point for a comprehensive categorization.

1. **Result of a measurement/outcome of a test:**
   e.g. „women with normal O'Sullivan test"
2. **Singular transgression of a threshold: a measurement yields results over/under a specific threshold**
   e.g. „if they [the following targets] are exceeded"
3. **Transgression of a threshold by a small margin:**
   e.g. „the threshold is exceeded by less than 15 mg/dL"
4. **Long-term transgression of a threshold/occurrence of a specific phenomenon:**
   e.g. „if ketonuria is negative for two weeks"
5. **Several transgressions of a threshold/events in a specific time interval:**
   e.g. „if low glucose values are encountered more than once at the same time of day"
6. **Decision/Case distinction:**
   e.g. „if the patient was (not) compliant with the prescribed diet"
7. **Attribute/property/characteristic trait:**
   e.g. "those with prediabetes"
8. **Consequence of an action/context dependent concept:**
   e.g. „if post-dinner or fasting glucose values achieve the threshold for insulin treatment as a consequence of increasing the amount of carbohydrates" or "hypoglycemia remote from meal or snack time"
9. **Categorization of ordinal measurement results:**
   e.g. "the result of ketonuria could be positive (++), positive (+), negative (+-), negative (-), negative (--)"

Equivalent models for the declarative knowledge categories:

For the introduced categories we want to find (all) possible equivalent models. We define two models as equivalent if (and only if) they accept the same information as input and the input produces the same output.

This section will give examples for equivalent models to describe the categories from the last section. Obviously this list is not conclusive and due to specific differences in wording there might be further possibilities to model a specific situation.

Category a)

Modelling the result of a specific measurement (e.g. fasting blood glucose concentration, ketonuria). Depending on the measurement the outcome can be nominal (e.g. ketonuria positive) or a numeric value (e.g. fasting blood glucose concentration). In all these cases a raw concept will be used to model possible outcomes of the measurement followed by a

value abstraction concept covering the specific result. Raw numeric concepts can also be modelled in a specific unit (e.g. mg/dL).

Category b)

Modelling a singular transgression of a threshold. This can be realized by using a raw numeric concept to represent the result of the measurement. A value abstraction concept models the transgression of the threshold (true/false) depending on the value of the raw numeric concept.

Category c)

Modelling the transgression of a threshold by a small margin. There are several ways to model this situation, which differs slightly from category b).

1. A raw nominal concept represents the transgression of the threshold (true/false). A value abstraction concept is derived from the raw nominal concept when its value equals true.
2. Two raw nominal concepts model the transgression of the threshold (true/false) and the compliance of threshold+margin (true/false). Two value abstraction concepts are derived from the two raw nominal concepts when their value equals true. A logic abstraction concept is derived from the two value abstraction concepts connects them with AND. Finally, a value abstraction concept derived from the logic abstraction concept represents the whole phrase.
   Note: The two value abstraction concepts could be skipped, if it were possible to model inside the logic abstraction concept the combination (raw nominal no. 1 equals true AND raw nominal no. 2 equals true).

Category d)

Modelling the long-term transgression of a threshold or the occurrence of a specific phenomenon over a specific interval. First a raw numeric or nominal concept models the transgression of the threshold (true/false). A value abstraction concept is derived from the raw concept (if the raw concept equals true). Finally a pattern concept is used to model the temporal constraint.

Category e)

Modelling several transgressions of a threshold or events in a specific time interval. There are several ways to model this situation.

1. A raw nominal concept models the transgression of the threshold (true/false). A value abstraction concept is derived from it, if the raw nominal concept equals true. To model the temporal constraint a pattern concept can be derived from the value abstraction concept.
2. A raw ordinal concept models the transgression of the threshold (e.g. high/low). A value abstraction concept can be derived, if the raw ordinal concept equals low (or high respectively). A pattern abstraction concept can be derived from the value abstraction concept and model the temporal or value constraints.

Category f)

Modelling a decision or a case distinction. For each possible choice or case a raw nominal concept (true/false) is created. Derived value abstraction concepts model each possibly taken decision or possible case (if raw nominal = true).

Category g)

Modelling an attribute, a property or a characteristic trait. This situation can be modelled in many different ways.

1. A raw nominal concept for the property of a specific person (true/false) and a value abstraction concept derived from the raw concept, if it equals true.
2. Two raw nominal concepts (one for the property, one for the person's group) with values true/false. A logic abstraction concept combines both raw concepts (AND). Finally a value abstraction concept can be derived from a logic abstraction concept, if it equals true.
3. Again, two raw nominal concepts (one for the property, one for the person's group) with values true/false are used. Then a context concept, derived from the person's group generated while the property equals true, models the phrase in total.

Category h)

Modelling the consequence of an action or a context dependent concept. This situation can also be modelled in different ways. Note: The causal relationship is not preserved in these models. However, often this is not necessary.

1. Two raw nominal concepts (true/false) model the consequence and the action or the context and the concept. A logic abstraction concept combines both (AND). A value abstraction concept can be derived, if the logic abstraction concept equals true.
2. Again, two raw nominal concepts (true/false) model the consequence and the action or the context and the concept. A context concept can be derived from the first raw nominal concept generated while the second raw nominal concept equals true.

Category i)

Modelling the categorization of ordinal measurement results. A raw nominal concept models the outcome of the measurement with the possible results as values. Value abstraction concepts are derived from these raw nominal concepts for each possible result. Logic abstraction concepts can be used to unite concepts which belong to the same class (e.g. ketonuria ++ and ketonuria + is considered as result "positive"). For each class a value abstraction concept can be derived from the corresponding logic abstraction concept.

Problems occurring while modelling medical declarative knowledge:

This section lists a few typical problems for modelling done by a computer program ("machine modelling") and for modelling done by a modeler in person ("manual modelling").

Machine modelling:

- Duplicate concepts: e.g. "insulin therapy" with ID= 90916 and ID=90922, both as raw nominal concept with values=true/false.

- Wrong concept type: e.g. “week” (ID=90964) was modelled as a raw nominal concept with values true/false)
- Wrong labels: e.g. “their blood glucose concentration” (ID=90954) was labelled with [Disease]

Manual modelling:

As an engineer some medical phrases cannot be understood in a unique logic way. One example is the phrase “women […] who develop complications suggesting GDM (macrosomia, polihydramnios) [sic]”. From a strict logic point of view it is not clear whether the combination of macrosomia AND polyhydramnios or already one of these findings alone is an indication suggesting GDM. A similar example is the phrase “if they [the three targets: fasting BGC, 1h PPGC, 2h PPGC] are exceeded”. The context is fasting, 1hour and 2hour postprandial blood glucose which exceed specific thresholds. Again, it is not clear if all of the targets need to be exceeded to start action or if one transgression is already sufficient.

Another example are fuzzy thresholds such as “130-140 mg/dL” for the one-hour-postprandial glucose target. In a computer program it is necessary to fix the threshold at a specific value.

As a last example we mention medical terminology. For an engineer it is sometimes difficult to accurately understand the meaning of medical terms, e.g. are all of the following terms diabetes/diabetes mellitus/glucose intolerance referring to the same disease or are there differences in their meaning.

Thus it is particularly important to consult medical advice when modelling clinical guidelines with unclear passages.

Algorithm to analyze the structure of the models

To compare the structure of the manually and the automatically generated model an R script[2] was written. The package XML[3] was used to import the XML formatted data into the R environment.

Note: Both approaches model declarative knowledge in the GDM guideline. However, this does not mean that both approaches model the same phrases. For the automatically generated models, the declarative knowledge was marked by the computer program LASSIE. This knowledge does not necessarily correspond to the declarative knowledge considered as “Gold standard”, which was used for the manual modelling.

Graph theoretical background:

Nodes and edges:

The concepts in our XML document are partly derived from each other. For example the concept “pregnant woman” is derived from the concept “pregnancy”. The concepts can be

[2] www.r-project.org
[3] cran.r-project.org/web/packages/XML/index.html

seen as nodes and the relations between the concepts as edges. They are directed acyclic graphs. In terms of graph theory our model consists of many trees. The whole model of the clinical guideline is the union of many such trees, i.e. related concepts.

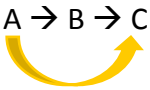
Transitive closure of a graph:

The transitive closure of a graph is a graph which contains an edge $\{x_i,x_j\}$ whenever there is a directed path from x_i to x_j in the original graph.

An example: A influences concept B and B influences concept C.

$$A \rightarrow B \rightarrow C$$

Thus, A has an (indirect) influence on concept C and we add an edge from A to C (yellow arrow).



Before the transitive closure there was only one concept with direct influence on C. This was element B.

After the transitive closure there are two concepts having either direct or indirect influence on C. Both elements A (indirectly through its influence on B) and B (directly) influence element C.

For each directed graph there is a unique adjacency matrix (see Table 1) representing the information which nodes are connected by an edge. There is an isomorphism between nxn-matrices with entries {0,1} and directed graphs with n nodes.

Table 1. Adjacency matrix for A, B, and C.

$x_{i,j}$	A	B	C
A	0	1	0
B	0	0	1
C	0	0	0

Each element $x_{i,j}$ can either be 0 or 1. 0 means there is no edge from x_i to x_j. Since there is no edge from C to A, the green field contains 0. Since there is an edge from B to C the blue field contains 1.

After the transitive closure, we have to add the edge from A to C to our matrix and thus insert the number 1 into the orange field (see Table 2).

Table 2. Adjacency matrix after the transitive closure.

x_(i,j)	A	B	C
A	0	1	1
B	0	0	1
C	0	0	0

Warshall algorithm:

The input of the Warshall algorithm is an adjacency matrix representing a graph. The output is the adjacency matrix of the transitive closure of this graph. The algorithm determines for each pair of nodes $\{x_i, x_j\}$ if there is a path between them. This path may lead over an arbitrary number of other nodes. The algorithm has a run time complexity of $O(n^3)$ and thus the same order as matrix multiplication.

Algorithm Warshall[4]
Input: Adjacency matrix *A* of relation *R* on a set of *n* elements
Output: Adjacency matrix *T* of the transitive closure of *R*.

Algorithm Body:
T := *A* [initialize *T* to *A*]
for *j* := 1 **to** *n*
    for *i* := 1 **to** *n*
        if *T_(i,j)* = 1 **then**
            a_i := *a_i* v *a_j* [form Boolean OR of row *i* and row *j*, store it in *a_i*]
    next *i*
next *j*
end Algorithm Warshall

Analysis of the models:

Network analysis:

To analyze the structure of the models we plot the concept nodes and the edges connecting them in a network graph. A grey arrow indicates a direct influence from one concept (shaft) to another concept (point). This means that the concept on the point of the arrow is derived from the concept on the arrow's shaft. Typically this derivative would be realized by a logic or value abstraction concept.

The numbers used in these visualizations are unique IDs assigned to each concept. They have no specific meaning and are only used to identify a concept. Note: The IDs in the

[4] See **http://www.dartmouth.edu/~matc/DiscreteMath/V.6.pdf**

automatically generated and the IDs in the manually created model are in no way related to each other.

The automatically generated model

Only very small clusters of concepts influencing each other occur. This corresponds to the result presented in the next section: Several concepts are duplicates and thus the structure is less nested.

We call a concept “isolated” when it does not interact with any other concept. This means it is neither derived from a concept nor does it influence any other concept. The graph of the automatically generated model shows 51 isolated concepts. They make up 38% of all concepts and correspond to the singular dots. Two examples (see Figure 6) are marked with a yellow circle.

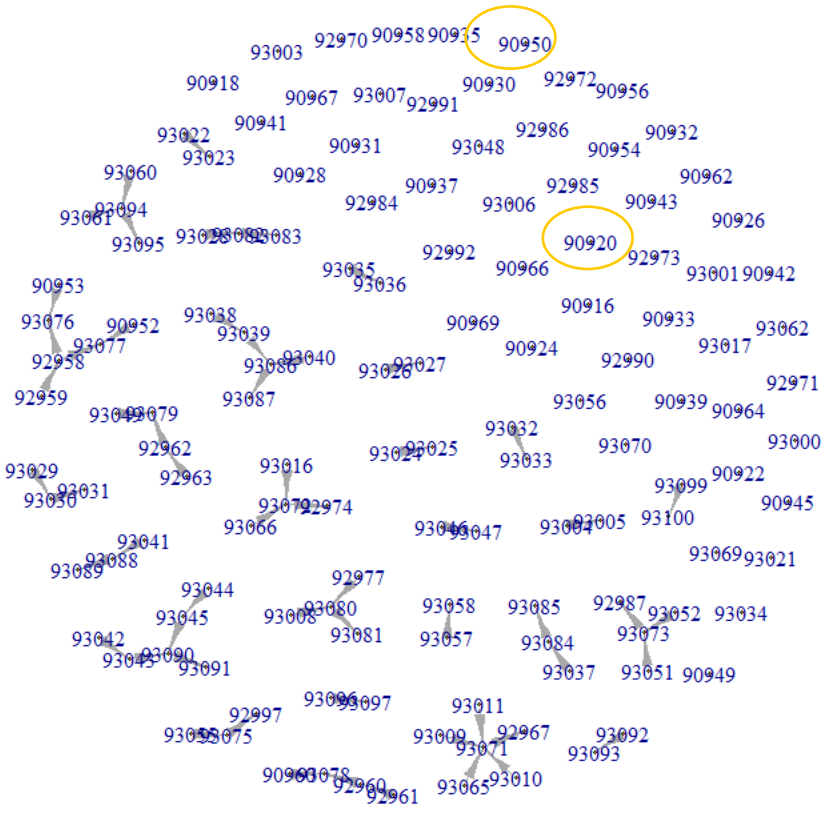


Figure 6. Graph representation of the automatically generated model.

The manually created model

However, in the manually generated model, there are far more and far bigger clusters.

Especially concepts such as sex (ID=100001) and its derivative woman (ID=100002) influence many other concepts (red circle and numerous grey arrows).

The only isolated concept without any interaction with other concepts is the raw nominal concept “low glucose values are encountered more than once at the same time of day”. It is marked with a yellow circle (see Figure 7).

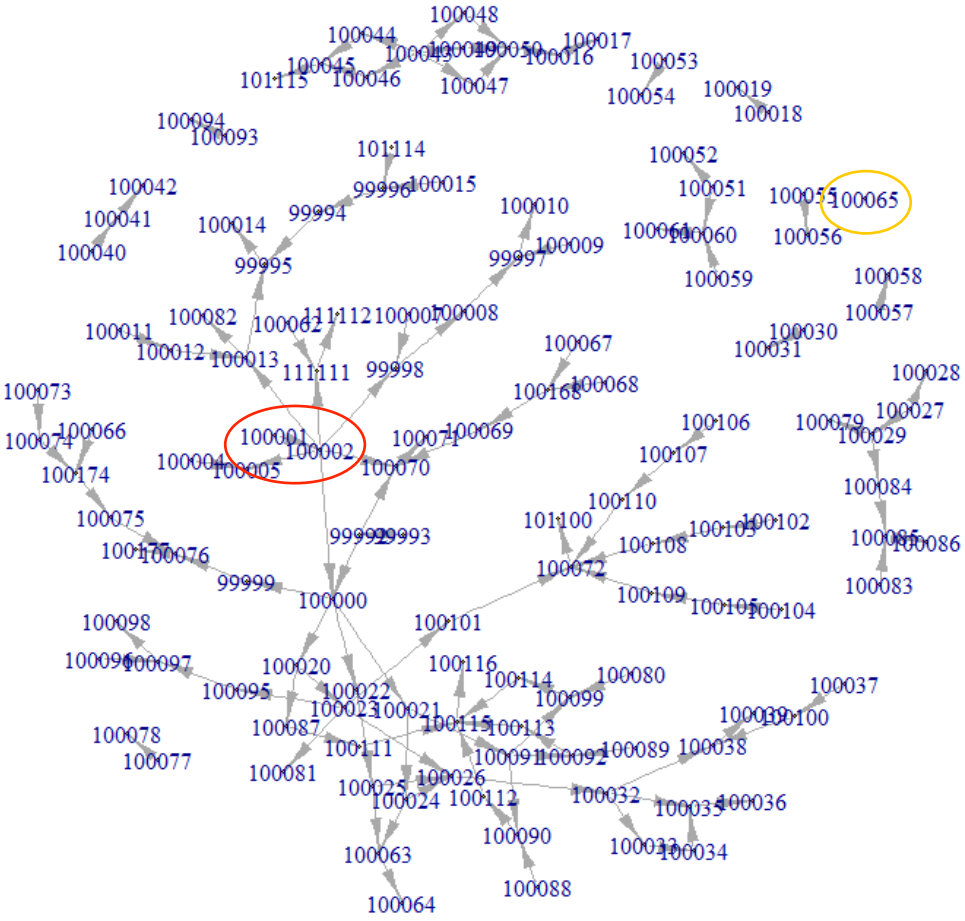


Figure 7. Graph representation of the manually generated model.

Detailed analysis of the automatically generated model:

Influence of one concept on other concepts: Before transitive closure

First we analyze the direct influence of one specific concept on a different concept before the transitive closure. Most of the concepts (58%) do not influence any other concept. Only 42% have an influence on exactly one different concept. The context concept women (derived from the concept sex C1522384) directly influences two other concepts. This is the maximum.

How many other concepts are directly influenced by one concept	**A → B, C, D, E,…** Note: percentages are rounded, thus the sum does not equal 100%
0	78 (58%)

1	56 (42%)
2	1 (1%)

Influence of one concept on other concepts: After transitive closure

After the transitive closure the situation does not change a lot. There are four concepts which influence either directly or indirectly three different concepts. This is the maximum obtained by the concepts sex (C1522384), fasting blood glucose (C0428568), insulin (C1579433) and the raw nominal concept "normoglycemia cannot be maintained". All 78 concepts without further influence naturally keep this state. Thus, 80% of the concepts influence at most one more concept directly or indirectly.

How many other concepts are directly or indirectly influenced by one concept	**A → B, C, D, E,…**
0	78 (58%)
1	29 (21%)
2	24 (18%)
3	4 (3%)

Influence of other concepts on one specific concept: Before transitive closure

Now we want to analyze by how many concepts a specific concept is directly influenced. The maximum is four obtained by the logic abstraction concept "(hypoglycemia, hyperbilirubinemia, hypocalcemia, erythremia)". This concept is derived from the concepts Hypoglycemia (C0020615), Hyperbilirubinaemia (C0020433), Hypocalcaemia (C0020598) and the raw nominal concept "erythremia".

How many other concepts directly influence a specific concept	**B, C, D, E,… → A**
0	90 (67%)
1	34 (25%)
2	10 (7%)
3	0 (0%)
4	1 (1%)

Influence of other concepts on one specific concept: After transitive closure

Now we will analyze by how many concepts a specific concept is directly or indirectly influenced. Two concepts reach the maximum of five other concepts: The value abstraction concept "metabolic complications" (Complications C0009566) was derived from the aforementioned logic abstraction concept "(hypoglycemia, hyperbilirubinemia, hypocalcemia, erythremia)" and is thus also influenced by the four predecessors. The second concept reaching the maximum was the logic abstraction concept "insulin is required because the fasting blood glucose concentration is high", which is derived from the value abstraction concepts "insulin is required" and "fasting blood glucose concentration is high".

How many other concepts directly influence a specific concept	**B, C, D, E,… → A**
0	90 (67%)
1	22 (16%)
2	9 (7%)
3	9 (7%)
4	3 (2%)
5	2 (1%)

Detailed analysis of the manually created model:

Influence of one concept on other concepts: Before transitive closure

Most of the concepts (about 90%) directly influence 0 to 1 other concepts (1 is the most common value obtained by 86 of 131 concepts), the maximum is 6 (obtained by the concept C0043210 woman).

How many other concepts are directly influenced by one concept	**A → B, C, D, E,…** Note: percentages are rounded, thus the sum does not equal 100%
0	30 (23%)
1	86 (66%)
2	9 (7%)
3	2 (2%)
4	2 (2%)
5	1 (2%)
6	1 (2%)

Considering the model after the transitive closure, we obtain the total number of influenced concepts. This means that direct and indirect influences are counted. Most of the concepts influence 0 to 10 other concepts. Note: The concept itself is not counted. The maximum is 53 (obtained by the concept sex C1522384) closely followed by 52 (obtained by the concept C0043210 woman). Three concepts reach values around 40:

- 42 (obtained by concept C0085207 Gestational Diabetes)
- 41 (immediate derivation of concept C0085207 Gestational Diabetes)
- 38 (derivation of concept C0085207 Gestational Diabetes and C0043210 woman).

There are 30 concepts, which do not influence any further concept. Thus, they stand for themselves. They can be isolated concepts which do not interact with any other concept such as the concept “low glucose values are encountered more than once at the same time of day”. However, all other concepts do have concepts as input and can thus be seen as the completed representation of a specific declarative knowledge string.

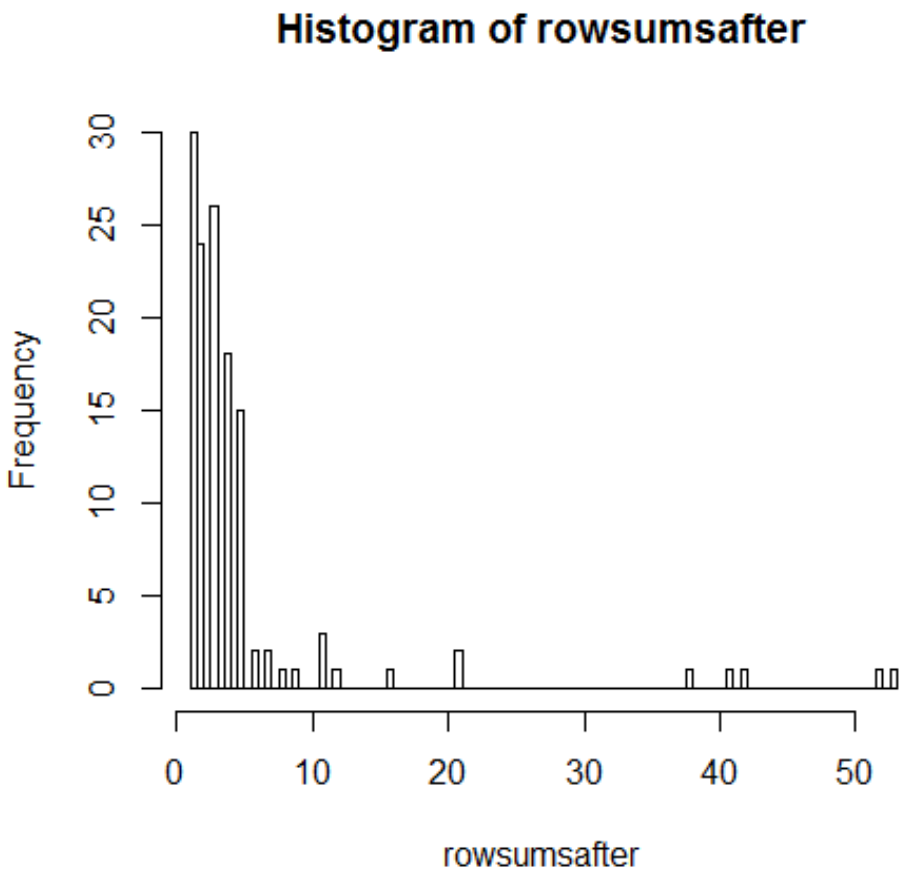


Figure 8. Histogram showing how many concepts are directly and indirectly (after transitive closure) influenced.

Influence of other concepts on one specific concept: After transitive closure

Now we want to analyze by how many concepts a specific concept is influenced. Looking at the histogram yields a maximum of 4 concepts which are used for the derivation of a new concept. This is the concept "if two or more elevated values are observed in a week period in the same interval", which is derived from four concepts representing an elevated value in a week in a specific interval (fasting, breakfast, lunch time, dinner).

How many other concepts directly influence a specific concept	**B, C, D, E,… → A** Note: percentages are rounded, thus the sum does not equal 100%
0	35 (28%)
1	70 (53%)
2	21 (15%)
3	3 (2%)

4	2 (2%)

And again, we have a look at the same data after the transitive closure. The maximum is 21 concepts which influence directly or indirectly the concept “if only two blood glucose targets are exceeded in one specific measurement [point...] and the threshold is exceeded by less than 15 mg/dL”. By transitive closure the number of concepts, which influence a specific concept can only rise. However, the 35 concepts which are influenced by no other concept directly, stay at the number of zero influences.

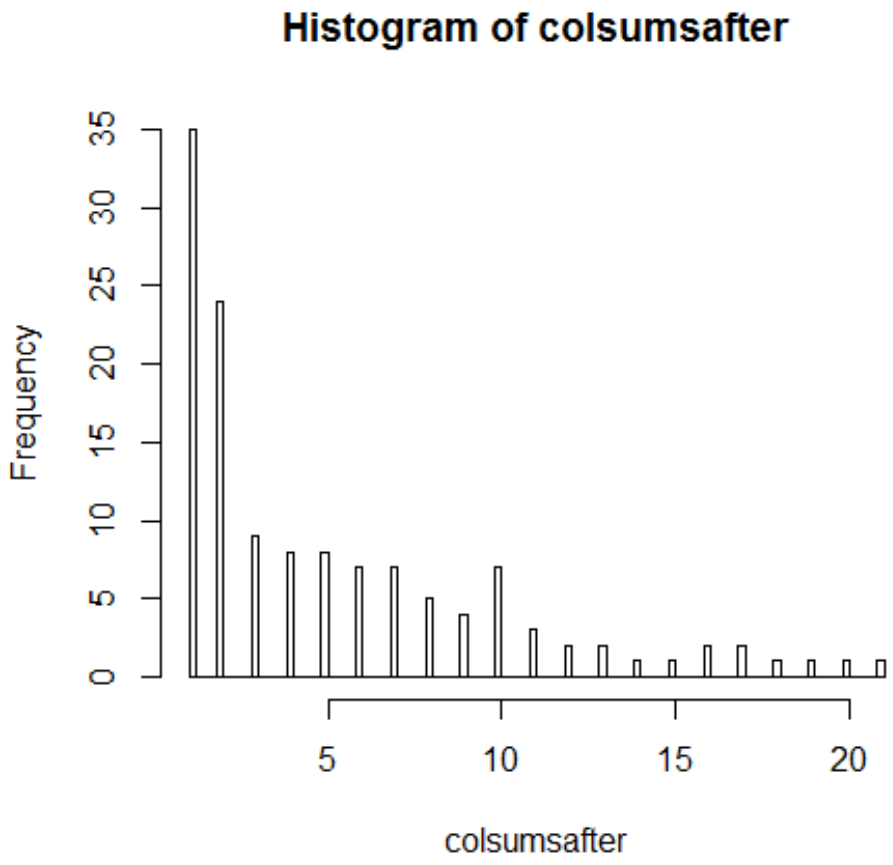


Figure 9. Histogram showing how many concepts influence one specific concept (including the concept itself).

Statistical analysis:

Concept types:

Descriptive statistical methods yield a comparison of the models’ structure and its trees. First we will compare the concept types used for modelling.

	Automatically generated model	Manually created model
Number of concepts	**135**	**131**
…thereof raw concepts	87 (65%)	41 (32%)
… … thereof raw nominal	79 (91%)	37 (90%)
… … thereof raw ordinal	2 (2%)	1 (3%)
… … thereof raw numeric	6 (7%)	3 (7%)
…thereof value abstraction concepts	28 (21%)	50 (38%)
…thereof pattern abstraction concepts	3 (2%)	13 (10%)
…thereof logic abstraction concepts	14 (10%)	22 (16%)
…thereof context concepts	3 (2%)	5 (4%)

Though the total number of concepts differs only slightly in machine vs. human modelling, the types of concepts vary hugely. From the two diagrams it becomes evident that the automatically generated model uses mostly raw nominal concepts, while the manually created model employs rather value abstraction and pattern abstraction models. This is due to the fact that during manual modelling we would rather derive concepts from already existing ones than define new primitive concepts. Thus, abstraction concepts appear more often than raw concepts in the second diagram. Nominal concepts were used only rarely in both methods. This is due to the fact that it is often not clear which nominal scale the guideline refers to, e.g. “high risk for GDM” might use a scale “low/high” or “low/medium/high” or another classification.

Interestingly, the ratio of raw nominal (91% vs. 90 %) and raw numeric concepts (7% vs. 7%) is almost the same for the automatically generated and the manually created models. Raw ordinal concepts were only very rarely used in both models (2% vs. 3%).

Logic abstraction concepts appear almost equally often (10% vs. 16%) though a bit more in the manually created model. Usually, logic concepts cannot be replaced by a different concept type. Since the manual model tends to use rather derived concepts, the number of logic concepts is slightly larger than in the automatically generated model.

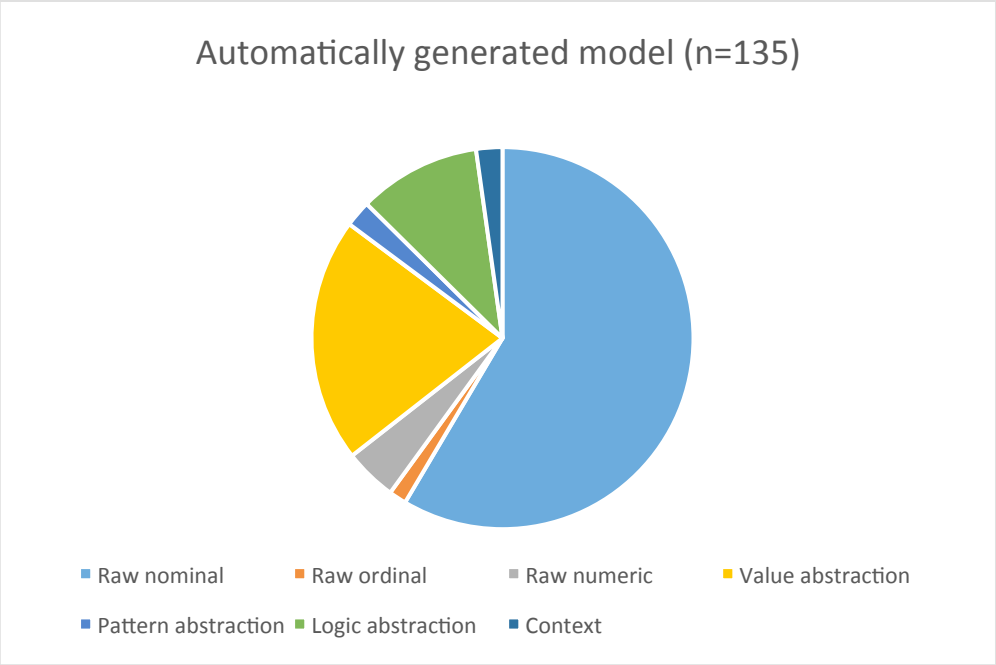


Figure 10. Distribution of the concept types in the automatically generated model.

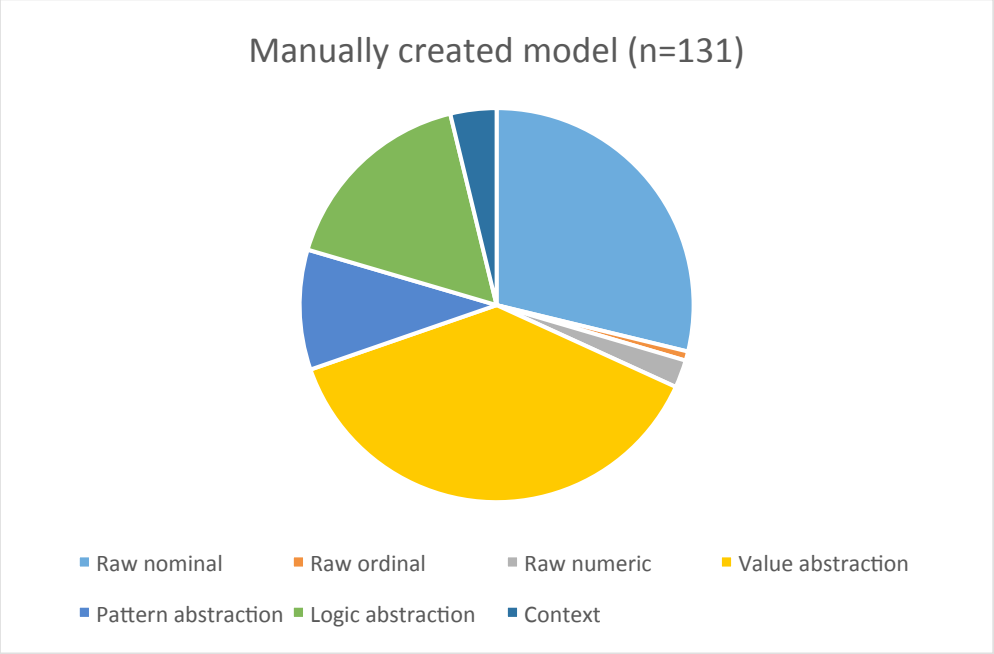


Figure 11. Distribution of the concept types in the manually generated model.

CUI (concept unique identifiers) concepts:

Concepts are important tools to classify and identify medical declarative knowledge. In this project, we used CUI (concept unique identifiers) from the UMLS to match the concepts used in the clinical guideline on GDM.

"Concept Unique Identifiers (CUI)

A concept is a meaning. A meaning can have many different names. A key goal of Metathesaurus construction is to understand the intended meaning of each name in each source vocabulary and to link all the names from all of the source vocabularies that mean the same thing (the synonyms)."

-- UMLS[5]

The automatically generated model

In the automatically generated model there were 76 concepts identified, though some of them occurred more than once.

An example for a duplicate would be Diabetes (C0011849), which was modelled four times as different concepts:

- Diabetes: Raw nominal concept (suspected)
- Subsequent risk for developing overt diabetes: Raw nominal concept (true/false)
- Long-term risk of diabetes: Raw nominal (true/false)
- History of diabetes: Raw nominal (true/false)

In total there were 44 different concepts left after deleting duplicates. Thereof were many diseases or syndromes such as Ketonuria (C0162275), GDM (C0085207) or Diabetes (C0011849). Additionally, therapeutic procedures such as Medical Nutritional Therapy (C0918259) were recognized.

Some concepts were also matched incorrectly. "Anti-hyperglycemic agents" was matched with the concept anti-anxiety agents (C0040616), probably due to the search key "anti-agents".

The manually created model

The manual model was much stricter in concept recognition. While the automatic model also matched the concept Diabetes (C0011849) with the concept "subsequent risk for developing overt diabetes", the manual model considered "risk for developing overt diabetes" as a different concept than diabetes and did not assign this CUI.

In the manually created model there were 19 concepts identified, none of them duplicates. The identified concepts were mostly diseases or syndromes such as polyhydramnios (C0020224) or ketonuria (C0162275). But also concepts such as supper (C2699739) or dietary regime (C0419178).

[5] https://www.nlm.nih.gov/research/umls/new_users/online_learning/Meta_005.html

Overview of CUI concepts found by either or both models.

The following table gives an overview of the CUI concepts identified in either or both models. Eight concepts were identified by both models, the automatic and the manual one. They are marked in green.

The grey columns provide information from UMLS (concept name and type). The orange columns provide information from the automatically generated model (key and CUI). Unfortunately, it was not possible to list the concept types and concept names for this model due to the many duplicates. The blue columns contain information on the manually created model (CUI, key, concept type and concept name).

MCM… Manually created model

AGM… Automatically generated model

Note: Scroll down to see the full excel list.

CUI Concept	Type of conc	Key used in t	CUI of Autor	CUI of Manu	Key used in t	Type of conc	Name (MCM)
Spontaneou	Pathologic fu	Miscarriage	C0000786				
Autoantibod	Immunologi	Autoantibod	C0004358				
Body Mass In	Diagonistic P	Body Mass In	C0005893				
Cardiovascul	Disease or Sy	Cardiovascul	C0007222				
Complicatior	Pathologic fu	Complicatior	C0009566				
Obstetric De	Therapeutic	Obstetric De	C0011209				
Diabetes	Disease or Sy	Diabetes	C0011847				
Diabetes Me	Disease or Sy	Diabetes	C0011849	C0011849	diabetes me	raw nominal	overt diabete
Diabetes Me	Disease or Sy	Type 2 Diabe	C0011860				
Glucose	Organic cher	Glucose	C0017725				
Hemoglobin	Amino acid,	Glycated hae	C0017853				
Polyhydramn	Pathologic function			C0020224	polyhydramn	raw nominal	polihydramni
Hyperbilirub	Disease or Sy	Hyperbilirub	C0020433				
Hypocalcemi	Pathologic fu	Hypocalcaem	C0020598				
Hypoglycem	Disease or Sy	Hypoglycaem	C0020615	C0020615	hypoglycemi	raw nominal	hypoglycemia
Insulin	Amino acid,	Insulin	C0021641				
Insulin, isopl	Amino acid,	NPH Insulin	C0021658				
Ketones	Organic cher	Ketones	C0022634				
Monosaccha	Organic cher	simple sugar	C0026492				
Obesity	Disease or Sy	Obese	C0028754				
Oral Glucose	Diagonistic P	Oral Glucose	C0029161				
Pregnancy	Organism Function			C0032961	pregnancy	raw nominal	pregnancy
Anti-Anxiety	Pharmacolog	anti-agents	C0040616				
Underweigh	Finding	Underweigh	C0041667				
Woman	Population G	Woman	C0043210	C0043210	woman	raw nominal	woman
Gestational I	Disease or Sy	GDM, Gestat	C0085207	C0085207	Gestational I	raw nominal	GDM [raw]
Exceptionall	Finding	Macrosomia	C0158915	C0158915	Exceptionall	raw nominal	macrosomia
Ketonuria	Congenital a	Ketonuria	C0162275	C0162275	ketonuria	raw ordinal (	ketonuria
Plasma Gluc	Laboratory P	Glucose	C0202042				
Impaired glu	Disease or Sy	Impaired glu	C0271650				
Insulin Lispr	Amino acid,	Insulin Lispr	C0293359				
Intermediate	Amino acid,	Intermediate	C0304869				
Insulin, shor	Amino acid,	Insulin, Rapi	C0356365				
Prediabetes	Disease or Sy	Prediabetes	C0362046				
Fetal acidosi	Disease or Sy	Fetal acidosi	C0410959				
Dietary Regi	Therapeutic or Preventive Procedure			C0419178	Dietary Regi	raw nominal	dietary thera
Fasting bloo	Laboratory P	Fasting bloo	C0428568				
gram	Quantitative	gram	C0439208				
week	Temporal co	week	C0439230				
Lunch time	Temporal concept			C0585039	lunch time	raw nominal	luch time
Mealtimes	Temporal concept			C0587119	Mealtimes	raw nominal	meal/snackti
Glucose norr	Finding	normal gluc	C0860800				
Medical nutr	Therapeutic	Medical Nut	C0918259	C0918259	medical nutr	raw nominal	medical nutri
MILK FAT	Organic cher	fat milk	C0991850				
Impaired fas	Finding	Impaired fas	C1272092				
Body mass in	Finding			C1319441	Body mass in	raw nominal	severely obe
Carbohydrat	Disease or Syndrome			C1401162	carbohydrate	raw nominal	carbohydrate
Pregnancy; p	Disease or Syndrome			C1411686	pregnancy; p	raw nominal	prediabetes
sex	Organism att	Sex	C1522384	C1522384	sex	raw nominal	sex
Insulin [EPC]	Pharmacolog	Insulin	C1579433				
No ketonuria	Finding			C1850724	no ketonuria	value abstra	ketonuria ne

Concept names:

This section lists all concepts and their types no matter whether they could be matched with a CUI concept or not (see Table 3 and Table 4).

Table 3. List of automatically generated concepts and their types.

1	blood glucose concentrations	RawOrdinal
2	blood glucose concentrations are high	ValueAbstraction
3	fasting blood glucose	RawOrdinal
4	fasting blood glucose concentration is high	ValueAbstraction
5	insulin	RawNominal
6	insulin is required	ValueAbstraction
7	medical nutritional therapy	RawNominal
8	fasting glucose	RawNominal
9	BMI	RawNominal
10	impaired fasting glucose	RawNominal
11	impaired glucose tolerance	RawNominal
12	a severely obese woman	RawNominal
13	GDM	RawNominal
14	ketonuria is negative for two weeks	Pattern
15	NO ketonuria	ValueAbstraction
16	ketonuria	RawNominal
17	GDM	RawNominal
18	overt diabetes is suspected	ValueAbstraction
19	diabetes	RawNominal
20	glucose < 130 to 140 mg/dL	ValueAbstraction
21	glucose	RawNumeric
22	normoglycemia cannot be maintained	ValueAbstraction
23	normoglycemia cannot be maintained	RawNominal
24	not follow recommendations	RawNominal
25	Underweight 40	ValueAbstraction
26	Underweight	RawNumeric
27	GDM	RawNominal
28	ketonuria 2	Pattern
29	ketonuria	RawNominal
30	are unable to take an oral glucose load	RawNominal
31	hypoglycemia	RawNominal
32	hyperbilirubinemia	RawNominal
33	hypocalcemia	RawNominal
34	GDM	RawNominal
35	medical nutritional therapy	RawNominal
36	medical nutritional therapy for more than 1 month	Pattern
37	macrosomia	RawNominal
38	unable to take an oral glucose load	RawNominal
39	obese	ValueAbstraction
40	risk of developing type 2 diabetes	RawNominal
41	NO history of GDM	ValueAbstraction

42	history of GDM	RawNominal
43	risk of developing GDM	RawNominal
44	subsequent risk for developing overt diabetes	RawNominal
45	Additional risk factors for impaired glucose tolerance	RawNominal
46	history of GDM	RawNominal
47	higher risk of developing cardiovascular disease	RawNominal
48	high risk for gestational diabetes	RawNominal
49	high risk for recurrent GDM	RawNominal
50	prediabetes	ValueAbstraction
51	risk of macrosomia	RawNominal
52	risk of fetal acidosis	RawNominal
53	Long-term risk of diabetes	RawNominal
54	history of GDM	RawNominal
55	sex	RawNominal
56	women	Context
57	sex	RawNominal
58	women	Context
59	metabolic complications	ValueAbstraction
60	risk of macrosomia increased	RawNominal
61	risk of miscarriage and congenital anomalies	RawNominal
62	develop complications suggesting GDM	ValueAbstraction
63	history of diabetes	RawNominal
64	high risk for GDM	RawNominal
65	women	Context
66	sex	RawNominal
67	She should also be given advice regarding contraception and the planning of future pregnancies	RawNominal
68	a 100 gr OGTT test	RawNominal
69	The patient is instructed to take this decision	RawNominal
70	postprandial blood glucose	RawNumeric
71	Glycated hemoglobin	RawNominal
72	Serial glucose	RawNominal
73	Serial glucose	RawNominal
74	their blood glucose concentration	RawNumeric
75	week	RawNominal
76	autoantibodies	RawNominal
77	delivery	RawNominal
78	urinary ketones	RawNominal
79	The patient measures ketonuria	RawNumeric
80	gram	RawNominal
81	simple sugars	RawNominal
82	insulin therapy	RawNominal
83	insulin therapy	RawNominal
84	anti-hyperglycemic agents	RawNominal
85	intermediate acting insulin	RawNominal
86	Additional doses of rapid acting insulin	RawNominal
87	an intermediate-acting insulin	RawNominal

88	the patient is recommended to measure ketonuria 2 AND If ketonuria is negative for two week	LogicAbstraction
89	NPH insulin	RawNominal
90	Women with normal O'Sullivan test but who develop complications suggesting GDM (macrosomia AND In thos	LogicAbstraction
91	NPH insulin	RawNominal
92	Women with normal O'Sullivan test but who develop complications suggesting GDM (macrosomia AND In thos	ValueAbstraction
93	the initial doses of insulin	RawNominal
94	the patient is recommended to measure ketonuria 2 AND If ketonuria is negative for two week	ValueAbstraction
95	an intermediate-acting insulin	RawNominal
96	insulin aspart or insulin lispro	RawNominal
97	insulin is required because the fasting blood glucose concentration is high	LogicAbstraction
98	non-fat milk	RawNominal
99	insulin is required because the fasting blood glucose concentration is high	ValueAbstraction
100	fasting glucose is elevated	LogicAbstraction
101	insulin dosage	RawNominal
102	fasting glucose is elevated	ValueAbstraction
103	insulin	RawNominal
104	prediabetes or overt diabetes	LogicAbstraction
105	Results should be recorded in a glucose log, along with dietary information	RawNominal
106	prediabetes or overt diabetes	ValueAbstraction
107	women with GDM	LogicAbstraction
108	a simple sugar	RawNominal
109	women with GDM	ValueAbstraction
110	women at high risk for gestational diabetes who are unable to take an oral glucose load	ValueAbstraction
111	women at high risk for gestational diabetes who are unable to take an oral glucose load	LogicAbstraction
112	women with GDM	ValueAbstraction
113	women with GDM	LogicAbstraction
114	not follow recommendations related to physical activity	ValueAbstraction
115	not follow recommendations related to physical activity	LogicAbstraction
116	normoglycemia cannot be maintained by medical nutritional therapy	ValueAbstraction
117	normoglycemia cannot be maintained by medical nutritional therapy	LogicAbstraction
118	(impaired glucose tolerance or impaired fasting glucose)	LogicAbstraction
119	(macrosomia, polihydramnios)	LogicAbstraction
120	(BMI ? 30 kg/m2)	LogicAbstraction
121	Serial glucose [women]	ValueAbstraction
122	Serial glucose [women]	ValueAbstraction
123	GDM [women]	ValueAbstraction
124	urinary ketones [women]	ValueAbstraction
125	erythremia	RawNominal

126	polihydramnios	RawNominal
127	a measure of blood glucose	RawNominal
128	overt diabetes mellitus should receive appropriate education and treatment	RawNominal
129	(hypoglycemia, hyperbilirubinemia, hypocalcemia, erythremia)	LogicAbstraction
130	previous GDM	RawNominal
131	oral glucose tolerance test	RawNumeric
132	oral glucose tolerance test 6 to 12	ValueAbstraction
133	prediabetes	RawNominal
134	overt diabetes	RawNominal
135	normal glucose tolerance	RawNominal

Table 4. List of the manually generated concepts and their types.

1	GDM [raw]	RawNominal
2	GDM	ValueAbstraction
3	Women with GDM	ValueAbstraction
4	Woman with GDM [Logic]	LogicAbstraction
5	sex	RawNominal
6	woman	ValueAbstraction
7	pregnancy	RawNominal
8	pregnant woman	Context
9	high risk for GDM	RawNominal
10	woman with high risk for GDM	ValueAbstraction
11	woman with high risk for GDM [Logic]	LogicAbstraction
12	unable to take an oral glucose load	RawNominal
13	woman with high risk for GDM who are unable to take an oral glucose load [Logic]	LogicAbstraction
14	woman with high risk for GDM who are unable to take an oral glucose load	ValueAbstraction
15	normal O'Sullivan test	RawNominal
16	normal O'Sullivan test	ValueAbstraction
17	woman with normal O'Sullivan test [Logic]	LogicAbstraction
18	woman with normal O'Sullivan test	ValueAbstraction
19	macrosomia	RawNominal
20	polihydramnios	RawNominal
21	complications suggesting GDM (macrosomia, polihydramnios) [Logic]	LogicAbstraction
22	complications suggesting GDM (macrosomia, polihydramnios)	ValueAbstraction
23	woman with normal O'Sullivan test but who develop complications suggesting GDM (macrosomia, polihydramnios)	LogicAbstraction
24	woman with normal O'Sullivan test but who develop complications suggesting GDM (macrosomia, polihydramnios)	ValueAbstraction
25	ketonuria negative	ValueAbstraction
26	ketonuria negative for two weeks	Pattern
27	number of ketonuria measurements with result positive is equal or higher than 3	Pattern
28	post-dinner glucose values achieve the threshold for insulin treatment as a consequence of increasing the amount of carbohydrates [raw]	RawNominal

29	fasting glucose values achieve the threshold for insulin treatment as a consequence of increasing the amount of carbohydrates [raw]	RawNominal
30	post-dinner or fasting glucose values achieve the threshold for insulin treatment as a consequence of increasing the amount of carbohydrates [logic]	LogicAbstraction
31	post-dinner or fasting glucose values achieve the threshold for insulin treatment as a consequence of increasing the amount of carbohydrates	ValueAbstraction
32	dinner time	RawNominal
33	post dinner glucose level is elevated [raw]	RawNominal
34	post dinner glucose level is elevated	ValueAbstraction
35	overt diabetes is suspected	RawNominal
36	when overt diabetes is suspected	ValueAbstraction
37	fasting blood glucose concentration	RawNumeric
38	1 hour postprandial glucose	RawNumeric
39	2 hour postprandial blood concentration	RawNumeric
40	fasting blood glucose concentration > 95 mg/dL (5.3 mmol/L)	ValueAbstraction
41	1h postprandial glucose >= 130 to 140 mg/dL (7.2 mmol/L)	ValueAbstraction
42	2h postprandial blood concentration > 120 mg/dL (6.7 mmol/L)	ValueAbstraction
43	if they [the three targets: fasting BGC, 1h PPGC, 2h PPGC]are exceeded	LogicAbstraction
44	they [the three targets: fasting BGC, 1h PPGC, 2h PPGC]are exceeded	ValueAbstraction
45	patient WAS/WASN'T compliant with the prescribed diet	RawNominal
46	patient was compliant with the prescribed diet	ValueAbstraction
47	patient was NOT compliant with the prescribed diet	ValueAbstraction
48	blood glucose is elevated as a result of non-compliance to nutritional prescription	RawNominal
49	patient was not compliant for three or more different meals in a period of one week	Pattern
50	patient was not compliant for three or more different meals in a period of one week, and blood glucose is elevated as a result of non-compliance to nutrial prescription [Logic]	LogicAbstraction
51	patient was not compliant for three or more different meals in a period of one week, and blood glucose is elevated as a result of non-compliance to nutrial prescription	ValueAbstraction
52	patient follows recommendations related to physical activity	RawNominal
53	patient does not follow recommendations related to physical activity	ValueAbstraction
54	dietary therapy [raw]	RawNominal
55	dietary therapy	ValueAbstraction
56	target glucose levels are exceeded despite dietary therapy	LogicAbstraction
57	when target glucose levels are exceeded despite dietary therapy	ValueAbstraction
58	medical nutritional therapy [raw]	RawNominal
59	medical nutriotional therapy	ValueAbstraction
60	normoglycemia cannot be maintained by medical nutritional therapy	LogicAbstraction
61	if normoglycemia cannot be maintained by medical nutritional therapy	ValueAbstraction
62	good glycemic control	RawNominal
63	good glycemic control accomplished	ValueAbstraction
64	when good glycemic control is accomplished with medical nutritional therapy for more than 1 month	Pattern
65	ketonuria	RawOrdinal

66	ketonuria positive [Logic]	LogicAbstraction
67	ketonuria negative [Logic]	LogicAbstraction
68	ketonuria plusplus	ValueAbstraction
69	ketonuria plus	ValueAbstraction
70	ketonuria plusminus	ValueAbstraction
71	ketonuria minus	ValueAbstraction
72	ketonuria minusminus	ValueAbstraction
73	prediabetes	RawNominal
74	those with prediabetes	ValueAbstraction
75	previous GDM	RawNominal
76	all women with previous GDM	ValueAbstraction
77	overt diabetes mellitus	RawNominal
78	woman who has overt diabetes mellitus	ValueAbstraction
79	normal glucose tolerance	RawNominal
80	women with normal glucose tolerance	ValueAbstraction
81	overt diabetes	RawNominal
82	prediabetes or overt diabetes	LogicAbstraction
83	women with prediabetes or overt diabetes	ValueAbstraction
84	severely obese	RawNominal
85	a severely obese woman	LogicAbstraction
86	in a severely obese woman	ValueAbstraction
87	postprandial blood glucose concentrations are high [Logic]	LogicAbstraction
88	postprandial blood glucose concentrations are high	ValueAbstraction
89	low glucose values are encountered more than once at the same time of day	RawNominal
90	hypoglycemia	RawNominal
91	meal/snacktime	RawNominal
92	remote from meal/snack time	Context
93	hypoglycemia remote from meal/snack time [Logic]	LogicAbstraction
94	hypoglycemia remote from meal/snack time	ValueAbstraction
95	hypoglycemia remote from meal/snack time [... in] women with GDM [Logic]	LogicAbstraction
96	hypoglycemia remote from meal/snack time [... in] women with GDM	ValueAbstraction
97	patient acknowledges that she is not following nutritional prescription more than once [raw]	RawNominal
98	patient acknowledges that she is not following nutritional prescription more than once	ValueAbstraction
99	insulin is required because the fasting blood glucose concentration is high	RawNominal
100	medical contraindications to this level of physical activity	RawNominal
101	obstetrical contraindications to this level of physical activity	RawNominal
102	medical or obstetrical contraindications to this level of physical activity	LogicAbstraction
103	NO medical or obstetrical contraindications to this level of physical activity	ValueAbstraction
104	women with GDM and no medical or obstetrical contraindications to this level of physical activity [Logic]	LogicAbstraction
105	women with GDM and no medical or obstetrical contraindications to this level of physical activity	ValueAbstraction
106	fasting glucose exceeded but lower than threshold + 15 mg/dL	ValueAbstraction
107	breakfast time	RawNominal
108	lunch time	RawNominal

109	1h postprandial glucose exceeded but lower than threshold + 15 mg/dL	ValueAbstraction
110	1h pp breakfast glucose exceeded but lower than threshold + 15 mg/dL	Context
111	1h pp lunch glucose exceeded but lower than threshold + 15 mg/dL	Context
112	1h pp dinner glucose exceeded but lower than threshold + 15 mg/dL	Context
113	only twice fasting glucose target exceeded but lower than threshold	Pattern
114	only twice postprandial breakfast glucose target exceeded but lower than threshold	Pattern
115	only twice postprandial lunch glucose target exceeded but lower than threshold	Pattern
116	only twice postprandial dinner glucose target exceeded but lower than threshold	Pattern
117	if only two blood glucose targets are exceeded in one specific measurement [point...] and the threshold is exceeded by less than 15 mg/dL [logic]	LogicAbstraction
118	if only two blood glucose targets are exceeded in one specific measurement [point...] and the threshold is exceeded by less than 15 mg/dL	ValueAbstraction
119	two or more fasting glucose values elevated in a week period	Pattern
120	postprandial breakfast values elevated [raw]	RawNominal
121	postprandial breakfast values elevated	ValueAbstraction
122	postprandial lunch values elevated [raw]	RawNominal
123	postprandial lunch values elevated	ValueAbstraction
124	postprandial dinner values elevated [raw]	RawNominal
125	postprandial dinner values elevated	ValueAbstraction
126	two or more postprandial breakfast values elevated in a week period	Pattern
127	two or more postprandial lunch values elevated in a week period	Pattern
128	two or more postprandial dinner values elevated in a week period	Pattern
129	if two or more elevated values are observed in a week period in the same interval [logic]	LogicAbstraction
130	if two or more elevated values are observed in a week period in the same interval	ValueAbstraction
131	situation is kept for more than 1 additional week	Pattern

Extra: Analyze LASSIE's modelling by using a marked guideline

In this section we used an already marked guideline as input for LASSIE. Thus, it is possible to analyze and compare LASSIE's modelling to the manually created model.

Network and statistical analysis:

First we have a look at the network graph.

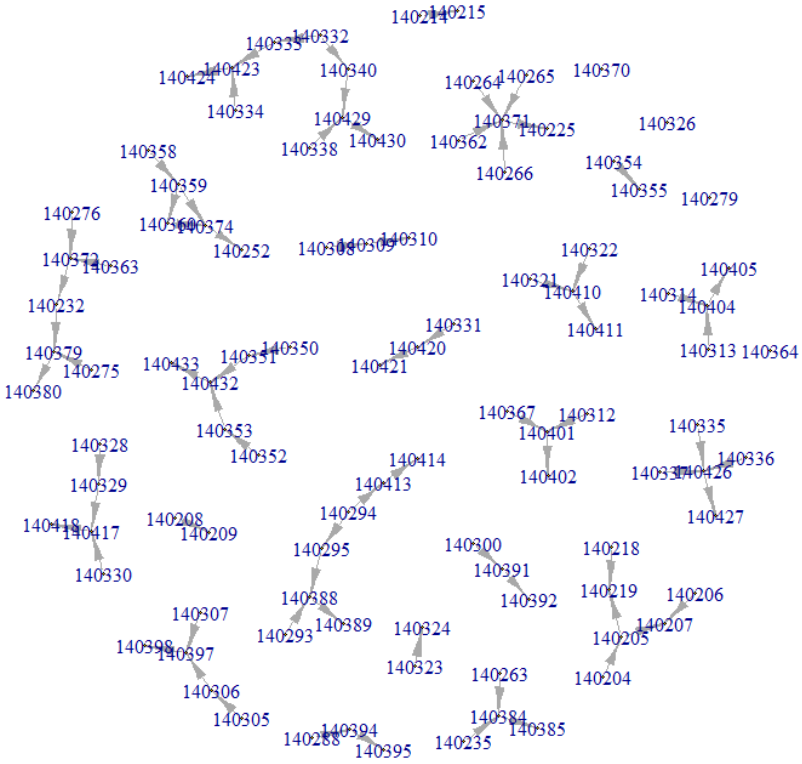


Figure 12. Network graph of the automatically generated model having the "condition" parts marked up before.

At the first glance we see that there are far less isolated concepts than in the LASSIE model presented before. There are also much more clusters. However, the number of concepts is much less (101 concepts) than before (135), also in comparison to the manual modelling (131).

This leads to the conclusion that LASSIE tends to identify more declarative knowledge than a modeler.

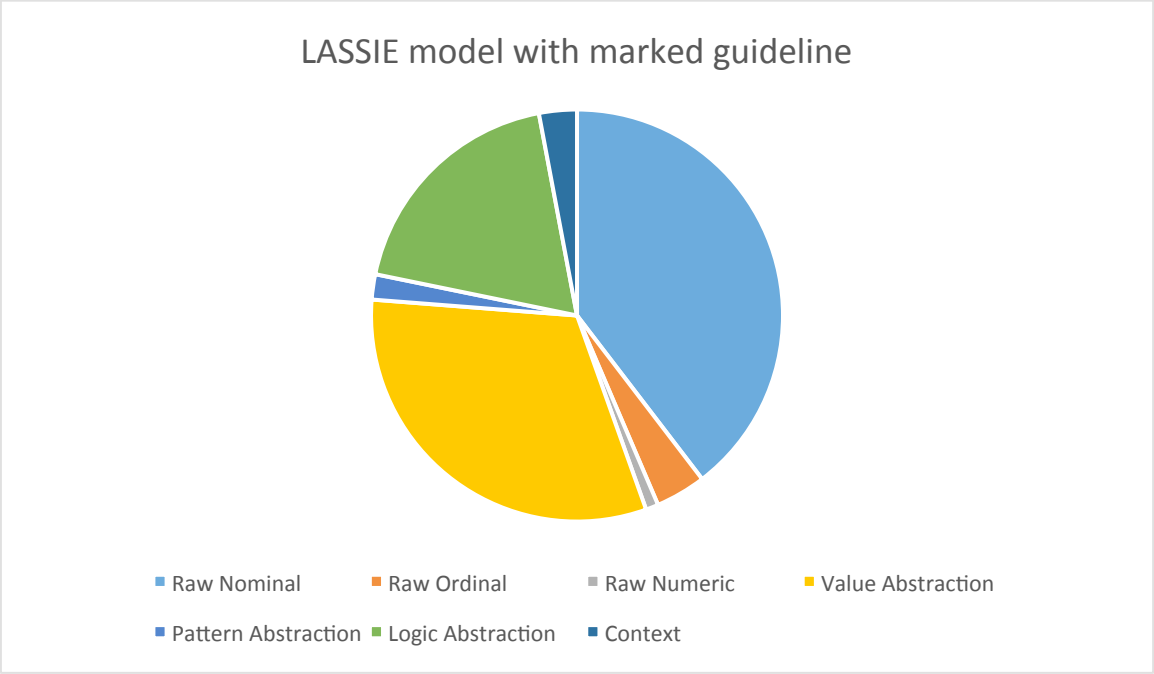


Figure 13. Distribution of concepts types in the automatically generated model from the marked-up guideline.

In the manual model the ratio of raw concepts is slightly lower (32% vs. 44%), on the contrary the ratio of pattern abstraction concepts is higher (10% vs. 2%). This shows that a modeler is more capable to derive from raw concepts instead of creating new ones. The ratio of value abstraction, logic abstraction and context concepts is almost the same in both models (see Table 5 and Figure 13).

Table 5. Concepts and their types in the automatically model from the marked-up guideline.

	LASSIE model with extra input	
Raw concepts	45 (44%)	
… thereof Raw nominal concepts		40 (88%)
… thereof Raw ordinal concepts		4 (9%)
… thereof Raw numeric concepts		1 (3%)
Value Abstraction	32 (32%)	
Pattern Abstraction	2 (2%)	
Logic Abstraction	19 (19%)	
Context	3 (3%)	
Total number of concepts	**101**	

Detailed analysis:

First we analyze the direct influence of one specific concept on a different concept.

Influence of one concept on other concepts: Before transitive closure

About one third of the concepts (29%) do not influence any other concept. About two third of the concepts (68%) have an influence on exactly one different concept. Four concepts each influence two more concepts. This is the maximum and obtained by blood glucose

(C0392201), medical nutritional therapy (C0918259), the value abstraction concept "BMI > 30 kg/m2" (sic) and the context concept "woman".

How many other concepts are directly influenced by one concept	**A → B, C, D, E,…** Note: percentages are rounded, thus the sum does not equal 100%
0	29 (29%)
1	68 (68%)
2	4 (4%)

Influence of one concept on other concepts: After transitive closure

The concept with the maximum influence on six other concepts was "blood glucose". "Medical nutritional therapy had an influence on five more concepts.

How many other concepts are directly or indirectly influenced by one concept	**A → B, C, D, E,…** Note: percentages are rounded, thus the sum does not equal 100%
0	29 (29%)
1	25 (25%)
2	35 (35%)
3	7 (7%)
4	3 (3%)
5	1 (1%)
6	1 (1%)

Influence of other concepts on one specific concept: Before transitive closure

Before the transitive closure a maximum of four concepts is used to derive a new concept. This was the logic abstraction concept "(hypoglycemia, hyperbilirubinemia, hypocalcemia, erythremia)". It was derived from the concepts "hypoglycemia", "hyperbilirubinemia", "hypocalcemia" and "erythremia".

The concept derived from three other concepts was the logic abstraction concept "ice cream, buns, cakes, packaged juice, plain sugar, or chocolate) or eating snacks not included in the prescribed diet)". It was derived from the concepts "Intracardiac echocardiography", "sugar" and "chocolate".

How many other concepts directly influence a specific concept	**B, C, D, E,… → A** Note: percentages are rounded, thus the sum does not equal 100%
0	45 (45%)
1	39 (39%)
2	15 (15%)
3	1 (1%)
4	1 (1%)

Influence of other concepts on one specific concept: After transitive closure

After the transitive closure the concept influenced by the most other concepts was the value abstraction concept “normal O'Sullivan test but who develop complications suggesting GDM (macrosomia, polihydramnios)”. It was influenced by a total of six other concepts. Three concepts were influenced by five other concepts: “metabolic complications”, “insulin is required because the fasting blood glucose concentration is high” and the logic abstraction concept “normal O'Sullivan test but who develop complications suggesting GDM (macrosomia, polihydramnios)”.

How many other concepts directly influence a specific concept	**B, C, D, E,… → A** Note: percentages are rounded, thus the sum does not equal 100%
0	45 (45%)
1	18 (18%)
2	11 (11%)
3	14 (14%)
4	9 (9%)
5	3 (3%)
6	1 (1%)

List of all CUI concepts:

This model contains 36 CUI concepts (cf. the following list). After removing the duplicates, 32 unique concepts remained. Thereof six concepts were identified manually and also by LASSIE (marked in blue):

1. Hypoglycaemia (C0020615)
2. GDM (C0085207)
3. Macrosomia (C0158915)
4. Ketonuria (C0162275)
5. Medical nutritional therapy (C0918259)
6. Sex (C1522384).

Table 6. List of unique concepts.

C1880312	Dietary Macronutrients Therapy
C0015663	Fasting
C1318607	Target HbA1c level
C0201802	Specific gravity measurement
C0994536	Intracardiac echocardiography
C0392201	blood glucose
C0523213	PAS
C0085207	GDM
C0021641	Insulin
C0939909	chocolate
C0242209	Sugar
C0595877	elevated blood glucose
C0421189	Good hypertension control
C0918259	Medical Nutritional Therapy
C0552468	Glucose mean value
C1511790	Detection
C0745343	insulin treatment
C0162275	Ketonuria
C0051767	DAS
C2825142	Result
C0005893	Body Mass Index
C0428568	Fasting blood glucose
C2585282	Blood glucose concentration
C0028754	Obese
C0009566	Complications
C0423548	Normal Rinne test
C0158915	Macrosomia
C0311468	Hyperbilirubinemia
C0020615	Hypoglycaemia
C0020598	Hypocalcaemia
C0421451	Patient date of birth
C1522384	Sex

List of all concepts and their types:

1	dietary therapy	RawNominal
2	fasting	RawNominal
3	target glucose levels	RawOrdinal
4	target glucose levels are exceeded	ValueAbstraction
5	specific measurement	RawNominal
6	ice	RawNominal
7	blood glucose	RawOrdinal
8	blood glucose targets are exceeded	ValueAbstraction
9	physical activity	RawNominal
10	normoglycemia cannot be maintained	RawNominal

11	not follow recommendations	RawNominal
12	GDM	RawNominal
13	normoglycemia cannot be maintained	ValueAbstraction
14	insulin is required	ValueAbstraction
15	insulin	RawNominal
16	not compliant	RawNominal
17	chocolate	RawNominal
18	plain sugar	RawNominal
19	blood glucose is elevated	ValueAbstraction
20	elevated blood glucose	RawNominal
21	GDM	RawNominal
22	good glycemic control	RawNominal
23	medical nutritional therapy	RawNominal
24	medical nutritional therapy for more than 1 month	Pattern
25	fasting glucose values	RawNominal
26	detection	RawNominal
27	insulin treatment	RawNominal
28	ketonuria	RawNominal
29	day	RawNominal
30	ketonuria is suspected	ValueAbstraction
31	NO ketonuria	ValueAbstraction
32	ketonuria	RawNominal
33	ketonuria is negative for two weeks	Pattern
34	ketonuria is suspected by the patient along the day	ValueAbstraction
35	ketonuria is suspected by the patient along the day	LogicAbstraction
36	women with GDM	LogicAbstraction
37	women with GDM	ValueAbstraction
38	elevated blood glucose is observed	ValueAbstraction
39	elevated blood glucose is observed	LogicAbstraction
40	good glycemic control is accomplished with medical nutritional therapy for more than 1 month	LogicAbstraction
41	good glycemic control is accomplished with medical nutritional therapy for more than 1 month	ValueAbstraction
42	women at high risk for gestational diabetes who are unable to take an oral glucose load	LogicAbstraction
43	women at high risk for gestational diabetes who are unable to take an oral glucose load	ValueAbstraction
44	normoglycemia cannot be maintained by medical nutritional therapy	ValueAbstraction
45	normoglycemia cannot be maintained by medical nutritional therapy	LogicAbstraction
46	not follow recommendations related to physical activity	ValueAbstraction
47	not follow recommendations related to physical activity	LogicAbstraction
48	fasting glucose values achieve the threshold for insulin treatment as a consequence of increasing the amount of carbohydrates	ValueAbstraction
49	fasting glucose values achieve the threshold for insulin treatment	LogicAbstraction

	as a consequence of increasing the amount of carbohydrates	
50	case of ketonuria detection	ValueAbstraction
51	case of ketonuria detection	LogicAbstraction
52	glucose results	RawNominal
53		RawNominal
54	BMI > 30 kg/m2	ValueAbstraction
55	erythremia	RawNominal
56	polihydramnios	RawNominal
57	BMI	RawNumeric
58	BMI > 30 kg/m2	ValueAbstraction
59	fasting blood glucose	RawOrdinal
60	fasting blood glucose concentration is high	ValueAbstraction
61	blood glucose concentrations	RawOrdinal
62	postprandial blood glucose concentrations are high	ValueAbstraction
63	normal O'Sullivan test but who develop complications suggesting GDM (macrosomia, polihydramnios)	ValueAbstraction
64	normal O'Sullivan test but who develop complications suggesting GDM (macrosomia, polihydramnios)	LogicAbstraction
65	OR	LogicAbstraction
66	OR	LogicAbstraction
67	OR	LogicAbstraction
68	gestation	RawNominal
69	obese	ValueAbstraction
70	high risk for gestational diabetes	RawNominal
71	develop complications suggesting GDM	ValueAbstraction
72	metabolic complications	ValueAbstraction
73	target glucose levels are exceeded despite dietary therapy	LogicAbstraction
74	target glucose levels are exceeded despite dietary therapy	ValueAbstraction
75	fasting, postprandial Breakfast, postprandial LUNCH, postprandial DINNER)	ValueAbstraction
76	fasting, postprandial Breakfast, postprandial LUNCH, postprandial DINNER)	LogicAbstraction
77	blood glucose targets are exceeded in one specific measurement point (fasting, 1-hour postprandial breakfast, 1-hour postprandial lunch, 1-hour postprandial dinner) and the threshold is exceeded by less than 15 mg/dL	LogicAbstraction
78	blood glucose targets are exceeded in one specific measurement point (fasting, 1-hour postprandial breakfast, 1-hour postprandial lunch, 1-hour postprandial dinner) and the threshold is exceeded by less than 15 mg/dL	ValueAbstraction
79	ice cream, buns, cakes, packaged juice, plain sugar, or chocolate) or eating snacks not included in the prescribed diet)	ValueAbstraction
80	ice cream, buns, cakes, packaged juice, plain sugar, or chocolate) or eating snacks not included in the prescribed diet)	LogicAbstraction
81	normal O'Sullivan test	RawNominal
82	not compliant for three or more different meals in a period of one	LogicAbstraction

	week, and blood glucose is elevated as a result of non- compliance to nutritional prescription	
83	macrosomia	RawNominal
84	not compliant for three or more different meals in a period of one week, and blood glucose is elevated as a result of non- compliance to nutritional prescription	ValueAbstraction
85	unable to take an oral glucose load	RawNominal
86	insulin is required because the fasting blood glucose concentration is high	LogicAbstraction
87	hyperbilirubinemia	RawNominal
88	hypoglycemia	RawNominal
89	insulin is required because the fasting blood glucose concentration is high	ValueAbstraction
90	hypocalcemia	RawNominal
91	are unable to take an oral glucose load	RawNominal
92	obese	RawNominal
93	obese woman	ValueAbstraction
94	DOB	RawNominal
95	more than 1	Context
96	at least the first week of follow-up	RawNominal
97	during at least the first week of follow-up	Context
98	woman	Context
99	sex	RawNominal
100	pregnant women	ValueAbstraction
101	pregnant	RawNominal

Conclusion:

Regarding conclusions one always needs to take into consideration that the two models model different declarative knowledge strings.

A striking difference is certainly the nested network of concepts in the manual model, whereas the computer model consists of an extremely high number of isolated concepts.

The absolute number of identified CUI concepts was much higher in the computer model even after duplicates were removed. Eight CUI concepts were identified by both models.

The computer model relied heavily on raw concepts, whereas the manual model preferred to work with a high number of value abstraction concepts. This observation corresponds to the results from the influence-analysis. In this analysis it was shown that the interactions between the concepts in the manual model are much higher than the ones in the computer model. Accordingly, the manual model used much more logic abstraction, pattern and context concepts.

An interesting detail is that the total number of concepts to be modelled from the guideline were almost the same (computer 135 vs. 131 manual). Also the ratio of the raw nominal, the

raw ordinal and the raw numeric concept among all raw concepts was almost the same in both models.

The conclusion from the extra analysis of LASSIE's models of a previously marked guideline is mainly that the number of concepts is much lower than the manually created model (101 vs. 131). Besides, LASSIE has a higher ratio of raw concepts and a lower ratio of pattern abstraction concepts. This finding provides evidence that a modeler's ability to derive from a raw concept instead of creating a new one, is more accomplished. LASSIE was able to identify more CUI concepts (32 vs. 19). Since several duplicates occurred, it is recommended to check if LASSIE's identification was correct.

Outlook:

- An analysis whether the CUI concepts identified in the computer model are actually correct is highly recommended.
- It would be interesting to analyze how the two modelling approaches cope with the "standard categories" of declarative knowledge strings listed above. Of particular interest would be which concept types are used in which situation. A comparison could start with a specific concept (CUI) or phrase and work out the structure of the model.
- The number of clusters in each model could be determined and investigated.
- It could be useful to merge the duplicates in the computer generated model to get a better understanding of the network graph structure. Without the duplicate concepts the table on CUI concepts could be extended to include concept types and concept names used by the computer model.

Appendix:

The clinical guideline on GDM with activities, conditions and definitions marked by LASSIE. All conditions were modelled as concepts in the manual model.

Gestational Diabetes Guideline (CSPT)

Introduction

Pregnancy is characterized by insulin resistance and hyperinsulinemia, thus it may predispose some women to develop diabetes. The resistance stems from placental secretion of diabetogenic hormones including growth hormone, corticotropin releasing hormone, placental lactogen, and progesterone, as well as increased maternal adipose deposition, decreased exercise, and increased caloric intake. These and other endocrinologic and metabolic changes ensure that the fetus has an ample supply of fuel and nutrients at all times. Gestational diabetes occurs when pancreatic function is not sufficient to overcome the insulin resistance created by changes in diabetogenic hormones during pregnancy.

Several adverse outcomes have been associated with diabetes during pregnancy. Importantly, the risk of these outcomes increases continuously as maternal fasting plasma glucose levels increase. However, there is no clear threshold that defines patients at increased risk. Adverse outcomes include: Preeclampsia

Hydramnios

Fetal macrosomia

Fetal organomegaly (hepatomegaly, cardiomegaly)

Birth trauma

Operative delivery

Perinatal mortality

Neonatal respiratory problems and metabolic complications (hypoglycemia, hyperbilirubinemia, hypocalcemia, erythremia)

If maternal hyperglycemia is present during organogenesis because of overt (also termed pregestational) diabetes, there is an increased risk of miscarriage and congenital anomalies.

Definition and diagnostic criteria

Gestational diabetes is defined as carbohydrate intolerance that begins or is first recognized during pregnancy. In 2010, the International Association of Diabetes and Pregnancy Study Group (IADPSG), recommended to classified it in two subtypes:

overt diabetes, which is defined with the same criteria than in general population but during the initial prenatal visit (Note: These patients will not be considered in MobiGuide): Fasting plasma glucose ? 126 mg/dL [7.0 mmol/L], or

A1C ? 6.5 percent using a standardized assay, or

Random plasma glucose ? 200 mg/dL [11.1 mmol/] that is subsequently confirmed by elevated fasting plasma glucose or A1C, as noted above

Gestational diabetes.

The diagnostic criteria for gestational diabetes have been a matter of debate since the publication of the IADPSG recommendations in 2010. The recommendation is based on outcome data reported in the HAPO study.

The Hyperglycemia and Adverse Pregnancy Outcome (HAPO) study included more than 23,000 pregnant women. After a 75-g oral two-hour glucose tolerance test, the risk of macrosomia increased as much as five-fold as fasting blood glucose concentration increased above 75 mg/dL (4.2 mmol/L), or one-hour glucose concentration increased above 105 mg/dL (5.8 mmol/L), or two-hour glucose concentration increased above 90 mg/dL (5.0 mmol/L), and the risk increased continuously across the spectrum of glucose results. There was also a positive, but weaker, correlation between increasing glucose concentration and maternal complications (eg, preeclampsia) and neonatal metabolic morbidity (eg, hypoglycemia, hyperbilirubinemia), but not long-term childhood morbidity, such as obesity at age two years in a small subset of offspring. Of note, women with significant hyperglycemia were excluded from the HAPO analysis (exclusion criteria: fasting glucose concentration greater than 105 mg/dL [5.8 mmol/L], two-hour glucose concentration greater than 200 mg/dL [11.1 mmol/L], or a random glucose concentration later in gestation greater than 160 mg/dL [8.9 mmol/L]).

On the light of the results of the HAPO study, the International Association of Diabetes and Pregnancy Study Groups (IADPSG) and the American Diabetes Association proposed criteria for Gestational Diabetes based on the results of the 75-gram Glucose tolerance test

Plasma glucose mg/dl mmol/l

Fasting ? 92 ? 5.1

One hour ? 180 ? 10.0

Two hours ? 153 ? 8.5

The thresholds represent the glucose values at which the odds of infant birth weight, cord C-peptide (proxy for fetal insulin level), and percent body fat >90 percentile were 1.75 times the estimated odds of these outcomes at mean glucose levels, based on fully adjusted logistic regression models. Compared to women in the HAPO study with all glucose values below the thresholds, women who exceeded one or more of these thresholds had a two-fold higher frequency of large for gestational age infants and preeclampsia, and >45 percent increase in preterm delivery and primary cesarean delivery. Using an odds ratio of 2 for the thresholds defined a population with further increased frequencies of these outcomes, but the increase was modest and resulted in failure to identify many women who were at almost comparable risk.

However, although the ADA has assumed the IADPSG recommendations, the American College of Obstetrics and gynecologists (ACOG) does not recommend this approach "because there is no evidence that diagnosis using these criteria leads to clinically significant

improvements in maternal or newborn outcomes and it would lead to a significant increase in health care costs".

The National Institute of Health is planning a Consensus Development Conference to determine the optimal approach to screening and diagnosis in the United States.

Thus, in this guideline we are going to maintain the recommendations for diagnosis made by the Spanish Group for the study of Diabetes and Pregnancy which agree with the recommendations from the ACOG:

All pregnant women should be screened for GDM using the O'Sullivan test, that consist on a 50-g, 1-hour loading test. First trimester (high risk for GDM) Age ? 35

Body Mass Index (BMI)= mass (Kg)/ (Height (m))^2 ? 30 kg/m2

Previous GDM confirmed or suspected on the light of the perinatal outcomes, i.e. macrosomia.

Familiar history of diabetes in first degree relatives.

Second trimester (24 - 28 weeks) All the pregnant women non previously diagnosed of GDM

Third trimester All the pregnant women non previously screened

Women with normal O'Sullivan test but who develop complications suggesting GDM (macrosomia, polihydramnios). In those cases it would be performed directly a 100 gr OGTT test

The diagnosis of GDM can be made based on the result of the 100-g, 3-hour oral glucose tolerance test, for which there is evidence that treatment improves outcome. The plasma levels designated by the National Diabetes Data Group are appropriate to use (see table). Status Plasma level

mg/dL mmol/L

Fasting 105 5.8

1 hour 190 10.6

2 hour 165 9.2

3 hour 145 8.0

A positive diagnosis requires that two or more thresholds be met or exceeded. If only one point is exceeded, then fasting and 1h postprandial glucose control are adviced. In case of normality of all the glucose results during at least the first week of follow-up, random and sporadic tests would be enough to guarantee an appropriate carbohydrate tolerance. The patient receives nutritional recommendations and exercise recommendations. Note: These patients will be enrolled in MobiGuide.

The highly concentrated hyperosmolar glucose solution can cause gastric irritation, delayed emptying, and gastrointestinal osmotic imbalance, leading to nausea and vomiting. In case of patients unable to tolerate oral hyperosmolar glucose, the recommended option is:

Serial glucose monitoring -- Periodic random fasting and one-hour postprandial blood glucose testing -- is a monitoring option for women at high risk for gestational diabetes who are unable to take an oral glucose load. HbA1c ? 6% could help to guarantee that this option is a reasonable way to start with. The patient receives nutritional recommendations and exercise recommendations. Note: These patients will be enrolled in MobiGuide.

Rationale for treatment

Identifying women with GDM is important because meta-analysis of randomized trials has shown that appropriate therapy can decrease maternal and fetal morbidity, particularly macrosomia. The value of treatment of GDM was best illustrated in two large trials that randomly assigned women with mild GDM to a regimen of either diet/blood glucose monitoring/insulin as needed or routine obstetrical care. One trial was performed in Australia and the other in the United States. Both trials observed that treatment resulted in a significant decrease in the prevalence of macrosomia (US trial: 6 versus 14 percent with usual care; Australian trial: 10 versus 21 percent with usual care), and the reduction in macrosomia was associated with a significant decrease in shoulder dystocia in one trial and the composite outcome of shoulder dystocia, nerve palsy, bone fracture, and death in the other. The rate of cesarean delivery was significantly reduced in the treatment group of one trial, and not increased in the other trial, although an increase is often observed in women with gestational diabetes. Treatment did not lead to a significant reduction in the prevalence of neonatal metabolic abnormalities (eg, hypoglycemia, hyperbilirubinemia) in either trial. Fetal/neonatal death was rare in all groups in both trials.

Both trials reported lower pregnancy weight gain in the treated group than in the control group. The US trial also reported a significant decrease in the rates of preeclampsia and gestational hypertension in the treatment group (8.6 versus 13.6 percent with usual care) and the Australian trial noted a significantly reduced rate of gestational hypertension (12 versus 18 percent with usual care).

Monitoring

Glucose monitoring

Women with GDM should measure their blood glucose concentration at least four times daily (fasting and one hour after the first bite of each meal) to determine whether hyperglycemia severe enough to increase fetal risk is occurring. Results should be recorded in a glucose log, along with dietary information. This facilitates recognition of glycemic patterns and helps immeasurably in interpreting results stored in the memory of modern meters. Multiple daily measurements allow recognition of women who should begin insulin therapy and appear to decrease the risk of macrosomia. The decision to start insulin therapy (or not to start it) is taken by the endocrinologist. Note: in some not very frequent cases, the insulin therapy is not started after considering the specific patient personal context (e.g. gestational age = 39 week and fetus percentile < 60, low family support, analphabetism and post-prandial BG < 160 mg/dl and fasting BG < 120 mg/dl).

Although there are no data on the duration of good control sufficient to reduce the frequency of self-monitoring or the appropriate frequency of testing in GDM that is well controlled (see section 5.1.1 Glucose target) with nutritional therapy, we recommend decreasing the frequency of glucose monitoring when good glycemic control is accomplished with medical nutritional therapy for more than 1 month . Decrease the frequency of glucose monitoring means to measure blood glucose (fasting and 3 postprandial measurements) twice a week instead of every day .

The advent of continuous glucose monitoring has the potential to allow determination of peak postprandial glucose levels; future research should determine whether using CGM in management of GDM will improve outcomes.

Glucose target

The ACOG recommends the following targets, with insulin therapy initiated if they are exceeded.

Fasting blood glucose concentration ? 95 mg/dL (5.3 mmol/L)

One hour -postprandial glucose < 130 to 140 mg/dL (7.2 to 7.8 mmol/L) or

Two hour -postprandial blood concentration ? 120 mg/dL (6.7 mmol/L)

In 'Hospital de Sabadell', we recommend the patients to measure postprandial blood glucose one hour after the first bite of each meal. For one-hour postprandial glucose, 140 mg/dL is considered the bound to detect anomalous glycemic control.

Little guidance is available as to what proportion of measurements exceeding these thresholds should trigger intervention. Insulin should be considered if elevated blood glucose is observed (see 5.3.1 Insulin).

Once insulin has been started, the glucose values considered as a goal of treatment will be those recommended by the 5th Workshop-Conference on Gestational Diabetes Mellitus:

Fasting blood glucose concentration < 95 mg/dL (5.3 mmol/L)

Sixty- 90 minutes -postprandial glucose ? 120 mg/dL (6.7 mmol/L)

Glycated hemoglobin

Glycated hemoglobin (A1C) may be a helpful test in assessing glycemic control during pregnancy, particullarly when overt diabetes is suspected. It should be noted that A1C values tend to be lower in pregnant compared to nonpregnant women because the average blood glucose concentration is about 20 percent lower in pregnant women and, in the first half of pregnancy, there is a rise in red cell mass and a slight decrease in red blood cell life span.

Ketonuria

Although there is conflicting evidence as to whether ketonuria is associated with an adverse effect on cognitive development of the fetus, we routinely monitor fasting urinary ketones in women with GDM. The patient measures ketonuria using urine strips. Monitor ketonuria routinely means to measure ketones in the urine every day at fasting conditions. If ketonuria is suspected by the patient along the day, she could decide to additionally measure ketonuria before lunch or dinner. The results of ketonuria could be: a) positive (++); b) positive (+); c) negative (+/-); d) negative (-); e) negative (--).

Ketonuria indicates that the person is in a catabolic state and is breaking down fat, and can occur in anyone who has a negative caloric balance. Pregnant women develop elevated ??-hydroxybutyrate levels more rapidly than nonpregnant individuals during a 12- to 18-hour fast, and it is not known whether such elevated levels have an adverse impact on fetal development. Several studies have suggested that early maternal malnutrition can affect neurobehavioral development in children of women with diabetes. In one such study, plasma beta-hydroxybutyrate levels independent of glucose levels had an adverse association with cognitive development in pregnancies in women with prepregnancy diabetes, GDM, and in normal pregnancies.

If ketonuria is negative for two weeks , the patient is recommended to measure ketonuria 2 or 3 times a week . In case of ketonuria detection (the number of ketonuria measurements with result "positive" is equal or higher than 3 in a period of time of one week):

If the patient was COMPLIANT with the prescribed diet , the nurse decides to increase the carbohydrates intake either at dinner or at bedtime: the amount of carbohydrates at dinner or at bedtime is increased by 1 unit (10 grams) . If post-dinner or fasting glucose values achieve the threshold for insulin treatment as a consequence of increasing the amount of carbohydrates , starting insulin would be the best option.

If the patient was NOT COMPLIANT with the prescribed diet , the nurse insists the patient on the importance of eating enough carbohydrates . If the situation is kept for more than 1 additional week , insulin therapy should be started .

Therapy

Nutrition

Patients with GDM should receive nutritional counseling upon diagnosis and be placed on an appropriate diet. The goals of medical nutritional therapy are to:

- Achieve normoglycemia (see ACOG recommendation in 4.1.1 Glucose target)
- Prevent ketosis
- Provide adequate weight gain
- Contribute to fetal well-being

There is scanty level evidence to support most aspects of the nutritional prescription for GDM .

In clinical practice, women often require 1600 to 2200 kcal per day.

Weight during pregnancy Caloric requirement (kcal/kg/day)

Underweight 40

Ideal body weight 30

Overweight 22 to 25

Morbidly obese 12 to 14

Once the caloric needs are calculated, carbohydrate intake needs to be distributed across meals and snacks to blunt postprandial hyperglycemia. Carbohydrate intake is limited to less than 55 percent of total calories and should be distributed in 3 main meals and 2-3 snacks. In order to avoid fasting ketonuria and facilitate post-dinner glucose control, a bed-time snack may be needed. As a general recommendation, the patients are informed that complex carbohydrates, such as those in starches and vegetables, are more nutrient dense and raise postprandial blood glucose concentrations less than simple sugars, which should be avoided.

Close follow-up is important to ensure nutritional adequacy. Individual assessment and self blood glucose monitoring are used to determine and modify specific nutrition/food recommendations. If the patient acknowledges that she is not following nutritional prescription more than once , she receives specific recommendations about the importance of following nutritional prescription.

If insulin therapy is added to nutrition therapy, a primary goal is to maintain carbohydrate consistency at meals and snacks to facilitate insulin adjustments. This decision is taken by nurse + physician.

Exercise

The value of exercise in women with GDM requires further exploration to determine the potential range of benefits. Nevertheless, based on the data available in pregnant and in

nonpregnant individuals, we recommend the regular practice (? 4days/week) of light or moderate exercise (16 - 28 METs hours per week) as part of the treatment plan for women with GDM and no medical or obstetrical contraindications to this level of physical activity.

If the patient does not follow recommendations related to physical activity , she is insisted by the nurse on the importance of following the recommended practice of physical activity.

This section of the general guideline will be further developed in more detail.

Pharmacologic therapy

If normoglycemia cannot be maintained by medical nutritional therapy, then anti-hyperglycemic agents should be initiated. There is only one option in pregnant patients who require medical therapy aimed at controlling blood glucose: insulin (and some insulin analogs), which is the only recommended approach in Spain.

Insulin

Women with GDM are placed on insulin therapy when target glucose levels are exceeded despite dietary therapy.

In the two randomized trials in which diagnosis and treatment of mild GDM improved outcomes, only 20 and 8 percent of women, respectively, required insulin, while 80 and 92 percent of women, respectively, were treated satisfactorily with diet.

Insulin therapy is started if two or more elevated values are observed in a week period in the same interval (fasting, postprandial Breakfast, postprandial LUNCH, postprandial DINNER). The endocrinologist is the person who takes the decision to start insulin therapy.

If only two blood glucose targets are exceeded in one specific measurement point (fasting, 1-hour postprandial breakfast, 1-hour postprandial lunch, 1-hour postprandial dinner) and the threshold is exceeded by less than 15 mg/dL , then two situations are considered (only for the first time these situations are detected):

If the patient was NOT COMPLIANT with the prescribed diet (diet intake higher than the recommended amount of carbohydrates or eating not recommended food (such as ice cream, buns, cakes, packaged juice, plain sugar, or chocolate) or eating snacks not included in the prescribed diet) then the nurse or the doctor motivate the patient to be compliant with the diet. If the patient was not compliant for three or more different meals in a period of one week , and blood glucose is elevated as a result of non- compliance to nutritional prescription, insulin therapy should be started.

If the patient was COMPLIANT with the prescribed diet then the nurse or the doctor can postpone starting insulin therapy by changing the nutritional prescription at lunch or dinner. E.g. two 1h post-dinner values of 142; 145 in a women eating more than or equal to 50 carbohydrate grams for dinner, with persistant negative fasting ketonuria. In this example

the reduction at dinner of 10 grams of carbohydrates could be enough for glucose normality restoration. See table for other situations:

Breakfast? Lunch Dinner

Diet 1600 kcal 60 g CARBS --? 50 g CARBS 50 g CARBS --? 40 g CARBS

Diet 1800 kcal 85 g CARBS --? 75 g CARBS 55 g CARBS --? 45 g CARBS

Diet 2000 kcal 85 g CARBS --? 75 g CARBS 75 g CARBS --? 65 g CARBS

Furthermore, the newborn weight prognosis based on ultrasoud measurements may be taken into account in the decision of starting insulin therapy; For fetus considered large for gestational age, one hour -postprandial glucose values above 120 mg/dL are enough for starting insulin. On the contrary, when the fetus is considered small for gestational age, 140 mg/dl should be the threshold value for considering insulin treatment. Definition of Small Gestational Age (percentile < 10) and Large Gestational Age (percentile > 90).

The dose and type of insulin used is calculated based upon the specific abnormality of blood glucose noted during monitoring. One principle we have found useful is to start with the simplest regimen and increase the complexity as needed to address the particular situation. If the post-dinner glucose level is elevated, then an injection of rapid acting insulin is given just prior to dinner. Additional doses of rapid acting insulin can be necessary to maintain euglycemia after breakfast or lunch. If fasting glucose is elevated, intermediate acting insulin can be given at bedtime, then a total of four injections per day are needed.

In summary:

If insulin is required because the fasting blood glucose concentration is high, an intermediate-acting insulin (NPH insulin), is given before bedtime; an initial dose of 0.1-0.15 unit/kg body weight is utilize.

If postprandial blood glucose concentrations are high, insulin aspart or insulin lispro is given before meals at an initial dose of 4-6 units.

In a severely obese woman, the initial doses of insulin may need to be increased to overcome the combined insulin resistance of pregnancy and obesity. This should be personalized by the endocrinologist according to each patient's characteristics.

Subsequent (weekly) adjustments in the various components of the insulin regimen are made based upon the corresponding glucose levels. Because any insulin regimen requires serial readjustment of dosage in response to specific fasting or postprandial glucose levels, the starting dose should be considered just that, a starting point. Adjustments in insulin dosage may be done to achieve the above mentioned objectives.

Hypoglycemia remote from meal or snack time is rare in women with GDM, and is treated by administering 10 to 20 g of a simple sugar like orange juice or non-fat milk. The patient is instructed to take this decision after getting a measure of blood glucose. If low glucose values are encountered more than once at the same time of day, insulin doses are adjusted downward accordingly.

Long-acting insulin analogs (insulin glargine, insulin detemir) have not been studied extensively in pregnancy. However, insulin detemir has been recently approved by the EMEA for the treatment of women with diabetes and pregnancy. In vitro perfusion studies have demonstrated that insulin glargine does not cross the placenta, however, concern about transplacental transfer of glargine in vivo remains. Based on available data, we prefer use of human NPH insulin as part of a multiple injection regimen in pregnant women with GDM.

Oral anti-hyperglycemic agents

A systematic review by the Johns Hopkins University Evidence-based Practice Center for the Agency for Healthcare Research and Quality presented evidence from randomized trials and observational studies that:

- maternal glucose levels did not differ substantially between gravidae treated with insulin versus those treated with oral glucose-lowering agents, and
- there was no consistent evidence of an increase in any adverse maternal or neonatal outcome with use of glyburide, acarbose, or metformin compared with use of insulin.

The ADA and ACOG do not endorse the use of oral anti-hyperglycemic agents during pregnancy and such therapy has not been approved by the Unites States Food and Drug Administration for treatment of GDM. We concur, but metformin may be a good complement to diet (plus insulin if required) in obese women. This should be personalized according to each patient's characteristics. However, metformin crosses the placenta and, in one study, cord arterial levels were twice as high as maternal venous levels.

Obstetric management

Obstetric management will be similar to the usually applied for non GDM women. However, it is advised to add an additional US study during 28-32 gestational week to detect macrosomia.

Peripartum management

Maternal hyperglycemia should be avoided during labor to reduce the risk of fetal acidosis and neonatal hypoglycemia. The risk of adverse neonatal metabolic outcomes (hypoglycemia, hyperbilirubinemia, hypocalcemia, erythremia) is related to both antepartum and intrapartum maternal hyperglycemia and appears to increase with the degree of maternal hyperglycemia. A protocol for intrapartum metabolic control is being applied which includes e.v. insulin perfusion/ dextrose serum. However, for almost all women with

gestational diabetes, no insulin treatment is needed during labour and we are going to review this part of the protocol in order to simplify it.

Future risks

Most women with GDM are normoglycemic after delivery. However, they are at high risk for recurrent GDM, prediabetes (impaired glucose tolerance or impaired fasting glucose), and overt diabetes over the subsequent five years .

Recurrence

One-third to two-thirds of women with GDM will have GDM in a subsequent pregnancy. Women who have a recurrence tend to be older, more parous, and have a greater increase in weight between their pregnancies than women without a recurrence. Higher infant birth weight in the index pregnancy and higher maternal prepregnancy weight have also been associated with recurrent GDM.

Long-term risk of diabetes

A history of GDM is predictive of an increased risk of developing type 2 diabetes, type 1 diabetes, and cardiovascular disease.

As many as 20 percent of women with GDM have impaired glucose tolerance during the early postpartum period.

A systematic review and meta-analysis found that women with GDM were at significantly higher risk of developing subsequent type 2 diabetes than women with normoglycemic pregnancies (RR 7.43, 95% CI 4.79-11.51; 20 cohort studies including 675,455 women of who 10,859 had type 2 diabetes). The relative risk was 4.69 within the first five years after delivery and 9.34 more than five years after delivery. In fact, 10 to 31 percent of parous nonpregnant women with diabetes have experienced a pregnancy complicated by GDM prior to their diagnosis.

Waist circumference and BMI are the strongest anthropometric measures associated with development of type 2 diabetes in women with GDM. Type 2 diabetes develops in 50 to 75 percent of obese (BMI ? 30 kg/m2) women with a history of GDM versus fewer than 25 percent of women with GDM who achieve normal body weight after delivery. Other major risk factors are gestational requirement for insulin and early gestational age at the time of diagnosis (ie, less than 24 weeks of gestation). Additional risk factors for impaired glucose tolerance and overt diabetes later in life include autoantibodies (eg, glutamic acid decarboxylase, insulinoma antigen-2), high fasting blood glucose concentrations during pregnancy and early postpartum, higher fasting blood glucose at diagnosis of GDM and high glucose levels in oral glucose tolerance testing, neonatal hypoglycemia, and GDM in more than one pregnancy. Parity, large birth weight, and diabetes in a first-degree relative are less correlated with later diabetes.

GDM is also a risk factor for the development of type 1 diabetes, particularly in populations with a high prevalence of this disorder. In our country there is no recent data but one study found that more than 13% of women with GDM have positivity to islet cell antibodies (ICA). Thus, if there is suspected the possibility of type 1 diabetes, determination of plasma autoimmune markers are mandatory (GAD65 Ab; IAA).

Women with GDM are at higher risk of developing cardiovascular disease and developing it at a younger age than women with no history of GDM.

Follow-up and prevention of type 2 diabetes

ACOG, the ADA and the Fifth International Workshop Conference on Gestational Diabetes recommend long-term follow-up of women with GDM:

All women with previous GDM should undergo an oral glucose tolerance test 6 to 12 weeks after delivery, using a two-hour 75 gram oral glucose tolerance test. An abnormal fasting blood glucose level is diagnostic (diabetes if ? 126 mg/dL, impaired fasting glucose (IFG) if 100 to 125 mg/dL); however, sensitivity for diagnosis of diabetes is low. Impaired glucose tolerance (IGT) is diagnosed if the two-hour value is 140 to 199 mg/dL. Collectively, IFG and IGT are known as "prediabetes."

Women with an abnormal oral glucose tolerance test are then classified as having prediabetes or overt diabetes mellitus

Those with prediabetes should be counseled about their subsequent risk for developing overt diabetes and referred for discussion of management options (eg, lifetime modification such as medical nutritional therapy, indications for metformin). They should try to achieve their ideal body weight through diet and exercise and, if possible, they should avoid drugs that may adversely affect glucose tolerance (eg, glucocorticoids). They should have yearly assessment of glycemic status.

A woman who has overt diabetes mellitus should receive appropriate education and treatment. She should also be given advice regarding contraception and the planning of future pregnancies.

Women with prediabetes or overt diabetes should be counseled regarding the importance of good metabolic control prior to any future pregnancies.

Women with normal glucose tolerance should be counseled regarding their risk of developing GDM in subsequent pregnancies and type 2 diabetes in the future. Lifestyle interventions (weight loss, exercise) are clearly beneficial for reducing the incidence of these disorders. Drug therapy (eg, metformin) also may have a role in preventing future type 2 diabetes.

References

Adapted from

Screening and Diagnosis of diabetes mellitus during pregnancy. Treatment and course of gestational diabetes mellitus. Up to Date. Authors: Donald R Coustan, MD; Lois Jovanovic, MD

Guia asistencial de diabetes mellitus y embarazo. Grupo Espanol de Diabetes y Embarazo. Av.Diabetol 2006;22:73-87

Screening and diagnosis of gestational diabetes mellitus. Obstet Gynecol 2011;118:751-2. Committee opinion on Obstetric Practice