

糖尿病实体标注和关系标注规范

一. 实体标注

1.1 疾病 (Disease)

在本规范中，疾病是一个宽泛的概念。疾病是在一定病因作用下自稳调节紊乱而发生地异常生命活动过程，并引发一系列代谢、功能、结构的变化，表现为症状、体征和行为的异常。

标注的原则：

(1) 疾病属于 ICD-10 分类标中的一种，或可以在百度百科、医学百科等确认为疾病。

(2) 疾病的中文名称、英文名称、英文简称是三个实体，需要分别标注。如果英文简称或者英文名称是作为疾病中文名称的说明出现的时候则是整体归为疾病，例如糖尿病 (DM)-整体为疾病。

(3) 疾病的类型、状态等标注为该病的分期分型(Class)。

(4) 泛化的疾病不进行标注，比如慢性病，并发症等泛化疾病不进行标注

标注样例：

1. 目前，2 型糖尿病及其并发症已经成为危害公众健康的主要疾病之一，控制血糖是延缓糖尿病进展及其并发症发生的重要措施之一
2. 这些研究终点的 HbA1c，控制在 7.5% 左右，那么年龄较大、糖尿病病程较长、部分已有心血管疾病(CVD) 或伴 CVD 极高危因素的糖尿病患者进一步降低血糖，对 CVD 的影响将会如何？
3. 研究表明，随着 HbA1 水平的升高，发生 CVD 的危险性相对增加，而且糖尿病病史越长的患者治疗后 CVD 危险性减低越小；

1.12 疾病的分期分型(Class)

标注原则：

疾病的类型、状态等标注为该病的分期分型(Class)。

标注样例：

2 型糖尿病作为整体作为疾病，嵌套标注为 2 型作为疾病的分期分型

1.2 病因 (Reason)

病因指外界客观存在的生物的、物理的、化学的、社会的等有害因素，或者人体本身的心理及遗传的缺陷，当其作用于人体，可以引起致病效应者，称为病因，或致病因素。

标注原则：

（1）病因指外界客观存在的生物的、物理的、化学的、社会的等有害因素，或者人体本身的心理及遗传的缺陷，当其作用于人体，可以引起致病效应者，称为病因，或致病因素。

标注原则：

病因主要是引起本疾病的风险因素（像年龄、体重、家族史、其他疾病等），而发病机制是说明风险因素是如何影响激素变化，进而引发疾病的。注意两者的区别。一般病因包括生物因素，理化因素，营养因素，遗传因素，先天因素，免疫因素，心理和社会因素等。

标注样例：

若**体重增加**，可能加重**胰岛素抵抗**，增加 CVD 风险。**体重增加-病因，胰岛素抵抗-发病机制**

1.3 发病机制 (Pathogenesis)

发病机制是指在身体受到内外环境的影响而发生了失衡的病理过程。之所以称为机制，就是每种疾病都有其一定的发病规律，并表现全身或者局部（系统、组织、器官、细胞等等）的病理反应。也就是说明风险因素是如何影响激素变化，进而引发疾病的。

标注原则：

发病机制是体内激素的变化过程注意和病因的区别

标注样例：

1 型糖尿病发病后数年，多数患者的 **β 细胞完全破坏**，**胰岛水平极低**，糖尿病的临床表现特别明显。

β 细胞完全破坏-发病机制

1.4 临床表现 (Symptom)

临床症状包括症状、体征，病人直接表现出来的和需要医生进行查体得出来的判断。本规范中，症状泛指由疾病导致的患者对机体生理功能异常的自身体验和感觉或者异常表现。其对应的类型主要有**症状**和**体征**，**症状**是患者描述的主观感

受（自述症状）、体征则是医生给病人检查时发现的具有诊断意义的**征候**。同时在本规范中症状还包含治疗后症状。

标注原则：

本规范的临床表现主要是患者的症状和体征

标注样例：

1. 患者自述最近食欲**不振、嗜睡、怕冷**。
2. 川崎病主要临床表现为**发热、唇充血皲裂、有猩红热样皮疹**

1.5 检查方法(Test)

检查方法指通过实验室技术、医疗仪器设备证实患者具有某种疾病或出现某种症状而采取的检查手段、过程或者检查项目，为临床诊断、治疗提供依据。本规范包括实验室检查方法，影像学检查方法，辅助试验，对于疾病有诊断及鉴别意义的项目等。

标注原则：

1. 需要标注具体的检查方式，如：MRI 检查、尿常规、腹腔穿刺。
2. 检查大类名称也进行标注：如实验室的检查、放射类检查、病理类检查等
3. 如果涉及动物试验，那么具体的检查方法也是进行标注的

标注样例：

1. 糖尿病检查可以检测可以通过**尿常规等实验室检查**等进行检测。

1.6 检查指标 (Test_Items)

检查指标是指通过实验室检查中的具体检查小项我们归为检查指标。

标注原则：

1. 需要标注的特指实验室检查中通过血液检查的具体指标或者检验小项，比如**血糖，餐后血糖，即时血糖**等
2. **实验室检查**整体还是归为检查方法。

1.7、检查指标值(Test_Value)

检查指标是指通过实验室检查或者医疗仪器设备等得出的具体结果。

标注原则：

1. 指标值包括数值结果的**符号以及单位**
2. 也包括文字包括**文字类描述**阴性，阳性，有无，增减，高低等。

标注样例：

1. 有低血糖症状并且血糖<3.3 mmol / L
2. HbA1C 控制于<6.5%

2.1 药物 (Drug)

药物是指用来预防、治疗及诊断疾病的物质。在理论上，药物是指凡能影响机体器官生理功能及细胞代谢活动的化学物质都属于药物的范畴。常见实体类型是类固醇、中药、化学药、生物药物、抗生素等。在标注中仅标注药物具体的名称和某类具体类别的药物

标注原则：

1. 只标注药物的具体名称
2. 标注某类具体类别的药物
3. 泛化的药物不进行标注（比如抗炎药，抗糖药）

标注样例

1. 在选择**胰岛素促泌剂**治疗时，可根据患者血糖谱选择不同类型**胰岛素促泌剂**（5级），（1）以餐后血糖（PPG）升高为主者，宜选择**格列吡嗪、格列奈类促泌剂**。
2. “改良剂型的**格列吡嗪控释片**”，这里标注为**格列吡嗪控释片**即可，修饰词不要进行标注。

2.2 用药频率(Frequency)

给药频次，即一段时间内用药的频率与次数。

标注原则：

1. 标注具体的频率即可

标注样例：

1. 使用格列吡嗪是**每天餐后一次**
2. 口服格列吡嗪控释片**一天2次**
3. 二甲双胍片：**q6h(六小时一次)**

常见临床的用药频率缩写：

qd(每日一次)

bid(每日两次)

tid(每日三次)

qid(每日)

qod(隔日一次)

qw(每周一次)

biw(每周两次)
tiw(每周三次)
qow(隔周一次)
2w(每两周一次)
3w(每三周一次)
4w(每四周一次)
q1/2h(30 分钟一次)
qh(每小时一次)
q2h(二小时一次)
q3h(三小时一次)
q4h(四小时一次)
q6h(六小时一次)
q8h(八小时一次)
q12h(12 小时一次)

2.3 用药剂量 (Amount)

药物的剂量是指给药时对机体产生一定反应的药量，通常是指防治疾病的用量。

标注原则：

1. 标注具体的数值和单位即可。

标注样例：

1. 比如“500mg/d”
2. 如格列吡嗪，1 片，餐后，口服，3/日；注射用胰岛素，10U，午餐前 15 分钟，肌肉注射，1/日；

2.4 用药方法 (Method)

目前临床上给药途径多样。 其中常见的给药途径有皮下注射、 静脉注射、 口服、 涂抹等。

标注原则：

1. 标注用药的具体方式

标注样例：

1. 胰岛素一般采用肌肉注射的方式
2. 每天餐后口服二甲双胍片 1 片

2.5 非药物治疗(Treatment)

治疗是为了解决疾病或者缓解症状而施加给患者的治疗程序，干预措施。本规范非药物治疗不包括手术以及药物治疗的方式。

标注原则：

1. 在医院环境下进行的非药物性治疗，包括放疗、中医治疗方法、医学营养治疗等，比如“推拿”、“按摩”、“戒烟”等。
2. 一些生活建议也作为非药物治疗的方式。

标注样例：

1. 医生建议进行强效治疗其中包括对患者的**饮食控制**以及**合理运动**

2.6 手术 (Operation)

手术是指医生用医疗器械对病人身体进行切除，缝合等治疗，以刀，剪，针等器械在人体局部进行操作来维持患者的健康，是外科的主要治疗方法。

标注原则：

标注时仅标注手术的具体方式即可，**泛化手术不需要标注，比如手术治疗**

标注样例：

1. 糖尿病的前沿技术是进行**胰岛细胞移植手术**来改善患者的胰岛情况。
2. 患者也可以**进行代谢手术**等治疗方式。

2.7 不良反应 (ADE)

用药或者手术后产生的不良反应，包括药物的不良反应以及手术后的不良反应。

标注原则

1. 对于是否是用药后或者手术后产生的不良反应要进行严格的区分。
2. 药物不良反应中的症状与疾病本身的症状不是完全一致的，比如糖尿病的症状有体重下降；口服药的不良反应有皮疹，二者不完全一致，所以需要标注

2. 标注样例：

1. 服用二甲双胍片容易出现**眩晕，低血糖症状**。根据语境这里标注为药物的**不良反应**即可。

3.1 部位 (Anatomy)

部位通常由多种组织构成能行使功能的结构单位。其对应的类别有身体部位、感觉器官、组织、细胞、胚胎结构等概念。

标注原则：

1. 这里的部位主要是指**疾病发病部位和累及部位**
2. 不需要对**非疾病的部位**进行标注
3. 如果疾病名里面有部位，需要进行嵌套标注，并关联关系

标注样例：

1. 糖尿病会引起一系列的病变，可能影响到的是**视网膜**等部位。
2. 糖尿病足-疾病，足-部位，并且标注糖尿病和部位的关系

3.2 程度 (Level)

程度 (Level)，包括病情严重程度，治疗后缓解程度等。

3.3 持续时间 (Duration)

持续时间 (Duration)，主要是指症状持续时间，因为症状持续时间对诊断病情是很关键的信息。也包括用药持续时间，其中用药的持续时间需要和具体药物进行关系关联。

标注原则：

1. 主要指症状的持续时间

标注样例：

1. “头晕**一周**”的“**一周**”。

二. 关系标注

实体关系标注在命名实体标注的基础上展开的，规范是以**疾病**和**药物名称**实体为中心，对于来源于的医学资源中出现的实体赋予预先定义好的关系类型。主要是拓展疾病和药物名称这两个实体的关系。

疾病关系

1. 检查方法 -> 疾病 (Test_Disease)
2. 临床表现 -> 疾病 (Symptom_Disease)
3. 非药治疗 -> 疾病 (Treatment_Disease)
4. 药品名称 -> 疾病 (Drug_Disease)
5. 部位 -> 疾病 (Anatomy_Disease)
6. 病因->疾病 (Reason_Disease)
7. 发病机制->疾病 (Pathogenesis_Disease)
8. 手术->疾病 (Operation_Disease)

9. 分期分型->疾病 (Class_Disease)

10. 检查指标->疾病 (Test_Items_Disease)

药物关系

1. 用药频率 -> 药品名称 (Frequency_Drug)

2. 持续时间 -> 药品名称 (Duration_Drug)

3. 用药剂量 -> 药品名称 (Amount_Drug)

4. 用药方法 -> 药品名称 (Method_Drug)

5. 不良反应 -> 药品名称 (ADE_Drug)

标注原则

1. 实体关系的抽取范围是基于一句完整的话

2. 实体属性同时属于疾病或者不良反应的时候，是以语境优先