

1 简介

1.1 协约 CONCORD-2 将会建立全球癌症生存监测，此附件提供了各登记处需要提供数据的详细说明。2012 年 9 月 20-21 日 Ireland-cork 举办了协约工作组会议，有来自 48 个国家的成员参会，会后对此附件进行了较大修改。

1.2 协约研究收集的数据将会送到伦敦卫生和热带医学学院。每个登记处提供最多 10 个数据文件（每种癌症一个），我们预计会收到来自于最多 200 个癌症登记处的最多 2000 份数据文件。为了更有效地管理这些文件，我们不得不严格规定数据结构。此附件列出了所有规则。

1.3 首先，我们规定 10 个**检索部位**-胃、结肠、直肠、肝脏、肺、乳腺（女）、子宫颈、卵巢和前列腺，以及白血病。

1.4 其次，为了简洁，我们将使用“**癌症**”指所有浸润性的、原发的恶性肿瘤，包括血液系统恶性肿瘤。

1.5 **检索癌症**包括：

- 发生在**检索部位**
- 在癌症登记处所覆盖地区的**常住居民**中诊断的
- 在登记处数据提交时间所覆盖**期间**内诊断的

1.6 大部分登记处将提供 10 种**检索癌症**的数据，但这并不是参加协约研究的要求。某些登记处只记录特定癌症（例如，乳腺、大肠、血液以及儿童癌症），某些登记处对于某些癌症没有登记患者的随访数据。

1.7 我们将主要分析 **1995 年 1 月 1 日到 2009 年 12 月 31 日**之间诊断的检索癌症患者的生存情况，希望各登记处尽可能多的提供此期间的数据，但是并不要求提供 1995-2009 年整个期间的数据，有些登记处是在 1995 年以后才开始运作。

1.8 然而，当您鉴定准备提交的数据时，应该确保（a）每年的**发病数据**是完整的，（b）所有患者生存状态的**随访数据**是完整的，至少访视到发病当年的 12 月 31 号，最好包括下一年度的访视结果。假如您有 2010 年以后的完整数据也可接受。例如：

- 例 1：发病数据 1995-2009，随访到 2009 年 12 月 31 日
- 例 2：发病数据 1998-2007，随访到 2008 年 12 月 31 日
- 例 3：发病数据 2000-2010，随访到 2011 年 12 月 31 日

1.9 某些参与此项目的**癌症登记处**覆盖整个国家的人口；其它的只覆盖国家的一部分（例如州、省、地区等）。

1.10 生存分析将主要针对**成年人**（定义为确诊年龄 15-99 岁）。也分析**儿童**（0-14 岁）急性淋巴细胞白血病的生存情况。

1.11 假如你们登记处收集良性（动态编码=0）、不确定（动态编码=1）或者原位癌（动态编码=2），例如宫颈原位癌，请在提交的数据中包含所有在**检索部位**诊断的肿瘤（动态编码 0-3）的记录。这使我们能够比较不同登记处诊断强度（见第三部分中的变量 22，动态）。然而，请注意，生存分析将仅限于**浸润性的，原发恶性肿瘤（动态编码 3）**。

1.12 在 1995-2009 年期间，大部分肿瘤登记处改用 ICD-O-3¹ 进行肿瘤部位、病理学和动态编码，取代原来的 ICD-9², ICD-10³ 或者 ICD-O-2⁴。2012 年 9 月在 Cork 的协约工作组会议上，各国一致同意使用 ICD-O-3。假如您的数据没有按照 ICD-O-3 编码，**请在提交数据前跟我们讨论**。

1.13 完整（包括年、月、日）的出生日期、确诊日期和死亡日期对于国际性生存分析十分重要。最近有一篇文章⁵解释了原理和证据，可以从 CONCORD 的网站上下载。在本附件的第三部分对相关变量进行了一个简要总结。假如您**不能**发送完整的出生日期、确诊日期和死亡日期，**请在准备与提交数据前与我们联系**。我们或许可以帮您获得伦理许可或者找到一个替代解决方案。

1.14 **所有的癌症数据文件都必须使用统一格式**。以下是数据文件中每个变量的详细信息。首先，请注意以下的普通规则，它们适用于几乎所有的变量：

- 所有的数据文件都会被检测是否严守方案。这是质量控制的第一步。数据提交后，方案遵守情况表将会发到各登记处。不符合方案的数据文件不能被使用，如果您对建立数据文件有疑问，**请在发送数据之前跟我们联系**。
- 每一条肿瘤记录的每一个变量都必须赋值（不能有缺失空项），包括**核心**和**可选**变量（见第三页）。
 - 如果您并没有收集某一变量的数据，请您务必在每条肿瘤记录中保留这个变量，不管它是核心变量还是可选变量，请不要让变量空白。例如，假如你没有收集种族（变量 8）信息，你应该在每条肿瘤记录的该变量赋值为 99（见 45-46 页的总结清单）。
 - 假如某条肿瘤记录中的某个变量不可用，请务必保留这个变量，无论它是核心变量或是可选变量。取决于每个变量有几位数（见 45-46 页），请用数字 9, 99 或 999 等代替这个变量的数值。例如，假如某条肿瘤记录缺失核心变量“性别”，你应该对此肿瘤记录的性别变量赋值 9。请不要用估算值替代。
 - 假如您选择不提供某个可选变量，请您务必在每条肿瘤记录中仍保留这个变量，并不要把

变量留为空白。例如，假如您不提供诊断时的分期（可选变量 23-37），请将其赋值为 9, 99 或者 999 等，具体根据每个变量（见 45-46 页）的位数而定。

- 假如您的肿瘤登记处日常工作中会用估算值替换缺失值（例如，月份），并且您提交的某些变量中包含估算值，**请告知我们：**我们将会向您索取一份估算程序的说明。
- 假如你们数据库的肿瘤记录中包含一个特殊代码（“flag”）来指明缺失值被估算，**请您在数据提交前跟我们讨论。**我们将会向您索取一份每个特殊代码的产生过程说明。
- 假如您已经因为登记处本地应用而通过添加特殊代码修改了标准编码框架（例如 ICD-O-3），请您在提交前**将数据重新编码**为标准格式。如果您对合适的转换方式有疑问，**请与我们联系。**
- 变量“**诊断时的分期**”是可选的，但是很多已收集了这些数据的登记处要求按照分期进行生存分析。我们只需要 2001 年及以后诊断病例的分期数据，对于在 1995-2000 年之间诊断的数据，肿瘤分期变量应当设为 9, 99 等等（见 45-46 页的清单）。2000 年及以前的分期数据不会包含在质量控制或生存分析中。

2 变量名和注释

2.1 以下是每条记录中将会用到的变量名顺序

核心变量

名称	简略说明
VAR1	国家
VAR2	登记处
VAR3	患者编码
VAR4	肿瘤编码
VAR5	IARC (国际癌症研究署) 检查标志*
VAR6	性别
VAR7	地区*
VAR8	民族/种族*
VAR9	出生日
VAR10	出生月
VAR11	出生年
VAR12	诊断日
VAR13	诊断月
VAR14	诊断年
VAR15	末次可知生命状态
VAR16	末次可知生命状态日
VAR17	末次可知生命状态月
VAR18	末次可知生命状态年
VAR19	诊断依据
VAR20	ICD-O-3 解剖位置
VAR21	ICD-O-3 形态学
VAR22	动态 (行为)

附加变量

名称	简略说明
VAR23	SEER Summary Stage 2000
VAR24	病理 T
VAR25	病理 N
VAR26	病理 M
VAR27	临床 T
VAR28	临床 N
VAR29	临床 M
VAR30	缩略 T
VAR31	缩略 N
VAR32	缩略 M
VAR33	Dukes' 分期
VAR34	FIGO 分期
VAR35	肿瘤大小
VAR36	检查的淋巴结数目
VAR37	受累淋巴结数目

*如果贵登记处不使用这些变量，见第二页 1.14

2.2 完整变量名、内容和编码细节在第三部分中给出。

2.3 缩略语表

缩写	英文全称	中文全称
AJCC	American Joint Committee on Cancer	美国肿瘤研究联合会
DCO	Death-certificate-only registration	仅依靠医学死亡证明书的登记
FIGO	International	国际妇产科联合会

Federation of Gynecology and Obstetrics		
IARC	International Agency for Research on Cancer, Lyon, France (WHO)	国际癌症研究机构，法国，里昂（WHO）
ICCC	International Classification of Childhood Cancers	国际儿童癌症分类
ICD	International Classification of Diseases (WHO)	国际疾病分类
ICD - O	International Classification of Diseases for Oncology (WHO)	国际疾病分类肿瘤学专辑
LSHTM	London School of Hygiene and Tropical Medicine	伦敦卫生与热带医学学院
NAACCR	North American Association of Central Cancer Registries	北美癌症登记联盟
SEER	Surveillance, Epidemiology and End Results programme (US National Cancer Institute)	监测、流行病学和最终结果项目
TNM	Tumour Nodes Metastasis (UICC)	
UICC	Union for International Cancer Control (Geneva)	国际癌症控制联盟

3 变量描述

变量 1 国家

四位数字变量，数值型

此变量包含 1 位大洲编码和联合国官方制定的 3 位国家编码。以下给出的这个 4 位国家编码必须包含在每条肿瘤记录中。

下表中只列举了中国相关的登记处编码。如需要其它国家的代码，请参考此附件的英语原版。

ASIA Eastern Asia

3156 中国

3344 中国香港

如果您的国家没有列出，请联系我们咨询。

以下的联合国网站中有 3 位国家编码，供参考，2012-10-31

<http://unstats.un.org/unsd/methods/m49/m49>

变量 2 登记处

1 至 3 位数变量，数值型（范围 1-950）

在您准备数据之前，我们将会提供给您这个编码，这个编码必须作为变量 2 包含在每条肿瘤记录中。

在生存分析时，，这个变量将跟国家编码（变量 1）一起被用来把数据文件和相应的寿命表关联起来。

变量 3 患者编码

最多 15 个字符的变量，纯数值或英文字母加数值型。

如果贵登记处的患者编码是纯数值型的，请不要在编码前加上“0”。例如，如果患者编码是九位数，请提交九个字符的编码（如 123456789），而不要用 000000123456789（15 个字符）。

这是贵登记处对每个登记的肿瘤患者给予的唯一编号。

患者编码可以是任何独特的字符串（数字或英文字母），但是请不要使用患者姓名、身份证号、社保卡号或者其他可以辨认的编码。**患者编码**必须包含在每条肿瘤记录中，如果我们在质量控制过程中发现可能的错误时，您可以根据编号检查这条记录。

同一个人在多条肿瘤记录中所使用的**患者编码**应当一致。患者编码与**肿瘤编码**（变量 4）一起，为研究中的每个肿瘤提供了唯一的标识码，便于质量控制，但同时不泄露患者的个人信息。

注意：可能有少数登记处在常规工作中不使用**患者编码**。这些登记处将需要为数据文件中的每一例患者编一个唯一码。这个编码将被用来识别多元原发肿瘤。假如我们在质量控制过程中发现了错误，登记处可以根据该唯一编码找到该患者的所有肿瘤。如有疑问，[请与我们联系](#)。

变量 4 肿瘤编码

不超过十个字符，纯数值或英文字母加数值型。

如果贵登记处的肿瘤编码是纯数值型的，请不要在编码前加上“0”。例如，如果肿瘤编码是六位数，请提交六个字符的编码（如 123456），而不要用 0000123456（10 个字符）。

这是贵**登记处**对登记的每一个肿瘤给予的编号。肿瘤编码与**患者编码**（变量 3）一起使得同一患者的多元原发肿瘤可以被识别。

主要的生存分析将包括 1995-2009 年（或者你们登记处所提供的数据期间，见 1.8）期间诊断的在**检索部位**的原发的、浸润性的恶性肿瘤。

假如贵登记处提交了 1995-2009 年确诊的所有患者数据，如果 2000 年一个患者被诊断为浸润性原发乳腺癌，随后又在 2005 年被诊断为另一种癌症-浸润性原发结肠癌，以上两种癌症的生存分析中都将包含此患者。

变量 5 IARC CHECK 检查标记

一位数字变量 数值型

我们将使用这个变量来避免发给您不必要的要求去核查那些已经核查并修正过的变量（检查标志=2,3,或 4）。

编码 意义

- 1 = 肿瘤未被 IARC CHECK 检查
- 2 = 肿瘤已被 IARC CHECK 检查，没有警告或者错误
- 3 = 肿瘤已被 IARC CHECK 检查，所有警告或错误都已被修正
- 4 = 肿瘤已被 IARC CHECK 检查，但未作任何修改，因为登记处已确认原始记录是正确的
- 9 = 此变量将不会提供

如果你不提供该变量，请将该变量的每一条记录编码为 9。

变量 6 性别

一位数字变量 数值型

编码 意义

1 = 男性

2 = 女性

9 = 性别不明或者未知

请不要在此变量的基础上从数据集中剔除任何记录，即使患者的性别未知。

变量 7 地区

五位数字变量 数值型

在某些情况下，可以对贵登记处覆盖下不同的地理区域进行分别的生存分析。例如，假如您的登记处覆盖全国，生存分析可以针对您国内的地区（例如：省、州等）。假如您的登记处覆盖一个省或者一个州，类似的分析可以针对您覆盖范围内更小的地区（例如县）。

假如您需要我们提供类似的分析，您将需要在每条肿瘤记录中都包含一个合适的地理编码。

同时您将需要提供地理编码赋值清单（例如：每个数值所代表省份的名称）。

例如，美国某州的区县地理编码：

Code	Meaning
21001	= Kentucky, Adair
21003	= Kentucky, Allen
21...	= Kentucky, ...
99999	= 地区未知，或者您不提供此变量

我们将需要为每个您想要得到单独生存分析的地理区域（省，州，等等）构建单独的寿命表：
见附件 6：寿命表。

假如您不选择提供此变量，请在每条肿瘤记录中对该变量赋值 99999。

变量 8 民族/种族

一位或两位数字变量，数值型

在某些情况下，可以在人群中分种族/族裔进行生存分析。某些国家的癌症登记处收集了种族和/或族裔（美国），种族（澳大利亚、以色列和新加坡），族裔（新西兰、英国）或者国籍（迪拜、科威特）。

与之相比，大部分欧洲登记处并没有记录种族或族裔的信息。在某些国家，这是法律不允许的。

每个登记处提供的这个变量中种族/族裔的分类都不相同。

如果您需要我们提供此类分析，您将需要在每条肿瘤记录中都包含一个合适的民族/种族编码。

您将需要提供一份该变量的编码清单，每个编码对应一个种族名称或者族群。

以下是美国的例子：

编码 意义

- 1 = 白人，西班牙裔
- 2 = 白人，非西班牙裔
- 3 = 白人，未知是否西班牙裔
- 4 = 黑人，西班牙裔
- 5 = 黑人，非西班牙裔
- 6 = 黑人，未知是否西班牙裔
- 7 = 亚洲或太平洋岛民，西班牙裔
- 8 = 亚洲或太平洋岛民，非西班牙裔
- 9 = 亚洲或太平洋岛屿裔，未知是否西班牙裔
- 10 = 美国印第安人/阿拉斯加原住民，西班牙裔
- 11 = 美国印第安人/阿拉斯加原住民，非西班牙裔
- 12 = 美国印第安人/阿拉斯加原住民，未知是否西班牙裔

13 =其他，未特指或未知的种族，西班牙裔

14 =其他，未特指或未知的种族，非西班牙裔

15 =其他，未特指或未知的种族，未知是否西班牙裔

99 =不详或缺失，或变量未提供

在我们与相关的登记处讨论后，其它的种族/族裔分组也可被使用，但这种分析的范围将会被我们能否为每个民族/族群构建独立的可靠的生命表所影响。

如果您确实需要按种族或族裔分类的生存分析，我们将也需要对每个种族/族裔构建合适的寿命表，见附件 6：寿命表。

假如贵登记处不收集关于民族/族群的数据，或您选择不提供此变量，请在每条肿瘤记录中**对该变量赋值 99**。

变量 9-11 出生日期

完整且精确的出生日期是很重要的，因为需要用它来计算确诊年龄。进行年龄专率生存分析和后续的年龄标准化生存分析时，需要用其确定（诊断时的）年龄组。该变量也被用来计算确切的死亡年龄，继而从寿命表中选择合适的参考死亡率来计算预期生存率。

只有少数癌症登记处没有记录完整出生日期，大部分登记处记录了完整日期，出于法律、伦理或者管理原因，某些登记处在向外界提供这些信息时可能会遇到一些法律、伦理或监管问题。

方案主体中提供了完整日期的重要性的解释。我们也已发表了一篇同行评议的文章⁵列出论点和相应的证据。这篇文章表明了当完整日期不能提供时，会在质量控制中遇到的问题，以及更重要的是在生存估计和生存比较中会带来的偏倚。这个结论是基于大数据集的敏感性分析得出的。这篇文章可以从 **CONCORD** 的网站上获得（使用您的账号和密码）。

数据准备和分析将在伦敦卫生和热带医学学校（LSHTM）进行。LSHTM 的肿瘤生存组已经获得英国相应机构的法律和伦理许可，以接受并分析有完整出生、诊断及死亡日期（日，月，年）的肿瘤记录（**附件 12.1：法律许可；附件 12.2：伦理许可**）。

变量 9 出生日

一位或两位数字变量，数值型

1-31 = 出生日

99 = 患者出生日不明

注意：假如所有患者的出生日都不能提供，请告知我们。

注意：关于日期估算，请见下面(变量 12-14)的评论。

变量 10 出生月

一位或两位数字变量，数值型

1-12 = 出生月

99 = 患者出生月不明

注意：假如所有患者的出生月都不能提供，请告知我们。

注意：关于日期估算，请见下面(变量 12-14)的评论。

变量 11 出生年

四位数字变量，数值型

YYYY = 出生年，从 1895（当他在 1995 年确诊时为 99 岁）至今

9999 = 患者出生年不明

变量 12-14 诊断日期

诊断日期是登记处用来统计癌症发病率和生存概率的日期。

因为诊断时间是生存分析的起始点，完整准确的诊断时间是非常重要的。

少数登记处仅仅记录诊断年和月，一些其它登记处仅仅记录了自 1995 年以后开始的完整诊断日期。

方案主体中给出了完整日期为何如此重要的解释。我们已经发表了一篇同行评议的文章列出论点和相应的证据。这篇文章表明了当完整日期不能提供时，在质量控制中会遇到的困难，以及更重要的是在生存估计和生存比较中会带来的偏倚。这个结论是基于大数据集的敏感性分析得出的。这篇文章可以从 CONCORD 的网站上获得（使用您的账号和密码）。

某些登记处会采集不止一种可能的诊断日期（例如入院日期，活检日期，手术日期等）。在提交数据到 CONCORD 之前，您需要完成一份关于你们登记处编码规范的问卷（附件 15：癌症登记处问卷）。其中问题包括您从两个及以上可能日期中选择诊断日期的方法。

我们恳请您在准备数据提交时不要对任何日期中的缺失部分进行估算。假如某些肿瘤诊断日期的日和/或月已经被估算了：

- 请提供给我们用来估算日期中月和日的规则
- 假如您已经常规性的添加一个标志来标明明日期中的日和/或月已经估算，请在提交数据前与我们协商

变量 12 诊断日

一位或两位数字变量，数值型

1-31 = 诊断日

99 = 患者诊断日不明

注意：假如因为法律或其它原因患者的诊断日期之诊断日不能提供，**请告知我们。**

变量 13 诊断月

一位或者两位数字变量，数值型

1-12 = 诊断月

99 = 患者诊断月不明

注意：假如因为法律或其它原因患者的日期之诊断月不能提供，**请告知我们。**

变量 14 诊断年

四位数字变量，数值型

YYYY = 诊断年，从 1995 年起

9999 = 患者诊断年份不明

变量 15 末次可知生命状态

一位数字变量，数值型

该变量是**癌症登记处**末次得知的患者**生命状态**的编码。末次访试生命状态的日期记录在变量 16-18 中。

编码 意义

- 1 = 存活
- 2 = 死亡
- 3 = 失访
- 9 = 生命状态不可知

生命状态的记录可以通过“**主动**”随访或者“**被动**”随访两种方法，而有些登记处也许会同时采用这两种方法，我们将在下文详细说明。您需要在向 **CONCORD** 提交数据文件之前**完成癌症登记处问卷（附件 15）**，告诉我们贵登记处的常规的做法。如果您不确定您登记处应用何种随访程序，**请您在准备数据前联系我们。**

“主动”随访

主动随访是指癌症登记处在常规工作中主动寻找每个癌症患者生命状态信息，来自于患者的医生、或者医院，甚至家访。

假如贵登记处使用这种方式随访，末次生存状态核查时状态（生或死）不明的患者请使用编码 3，这些患者属于失访，末次得知该患者存活的日期应当体现在变量 16-18 中。

明确已经搬迁的患者也应当使用编码 3，因为他们也属于失访：迁移日期应当赋予变量 16-18。

“被动”随访

被动随访是指登记处通过一种或多种可靠来源接收所有登记患者生命状态的的相关信息。不同登记处之间此种信息来源差别较大，可能包含社会安全与健康保险文件，或者一个地区或全国的死亡人员名单。信息可能来自于多种途径，例如通过电脑于登记处数据库相连，人工扫描死亡目录，或者提供登记处管辖范围内所有死亡的细节。

在国际存活比较中被动随访的主要特征是：

- 登记处使用这种方式更新其数据，以进行本地的生存分析
- 登记处通过这种程序能够可靠地得知所有登记患者的死亡，而不仅仅是那些在死亡登记中有癌症作为死因的死亡人员
- 除非从一条或多条途径得到患者死亡的信息，登记处能够可靠地假定患者仍然存活。

假如你们登记处使用被动随访的方式，当末次登记处和提供的生命状态记录链接时，尚未得知其死亡的患者将被认为是存活。因此这些病例应被编码为 1（存活）。

但是，假如有一些登记的患者不能够被任何被动随访的程序追查，他们的生命状态可能保持不可知：这时候他们的生命状态应该被编码为 9（不可知）。

当您提交数据时，我们将要求您告知我们您末次通过数据库链接或其他随访方式确认患者生命状态的日期。

变量 16-18 末次得知患者生命状态的日期

这是末次得知**患者生命状态**（变量 15）的日期。

假如患者死亡，末次得知生命状态的日期应当是死亡日期。

假如已知患者死亡（变量 15 被编码为 2），但是死亡日期未知，末次得知患者生命状态的日期应当被编码为 99,99,9999（见 45-46 页清单）。

假如患者迁移，末次得知生命状态的日期应当是迁移日期。

假如患者失访，末次得知生命状态的日期应当是失访日期。

假如患者被认为是存活，但**没有**迁移或者失访，末次得知生命状态的日期应为以下的一种：

- 所有患者随访已完成的最近一年的 12 月 31 日
- 登记处最近查核患者生命状态的日期，例如与患者医生联系或家访的日期（如果登记处实行主动随访），或者登记处与死亡登记数据库相连的日期（如果登记处实行被动随访）（见变量 15）
- 登记处从数据库为本研究提取数据文件的日期。

变量 16 末次得知生命状态日期-日

一位或两位数字变量，数值型

1-31 = 末次得知生命状态日期-日

99 = 该日期中的日未知

注意：假如某个您的患者末次得知生存状态日期的日不能提供，请告知我们。

注意：关于日期的估算，见以上条目（变量 12-14）。

变量 17 末次得知生命状态日期-月份

一位或两位数字变量，数值型

1-12 = 末次得知生命状态日期-月

99 = 该日期中的月份未知

注意：假如某个您的患者末次得知生存状态日期的月份不能提供，请告知我们。

注意：关于日期的估算，见以上条目（变量 12-14）

变量 17 末次得知生命状态日期-年

四位数字变量，数值型

YYYY = 末次得知生命状态日期-年

9999 = 日期中的年未知

变量 19 诊断依据

一位数字变量，数值型

该变量指明对癌症诊断是建立在何种程度的确定性之上。

编码 意义

非显微镜检验

1 = 仅有临床诊断

2 = 临床检查无组织学诊断（内窥镜检查未进行组织活检，或者影像学例如 X 线、超声、计算机断层扫描[CT]或者核磁共振[MRI]）

3 = 临床诊断，无其他说明（例如，不知道编码“1”“2”是否适用）

显微镜下检验

4 = 细胞学证实（包括血片检查白血病）

5 = 组织学证实（包括骨髓活检检查白血病）

6 = 显微镜下证实，无其他说明（例如，不知道编码“4”“5”是否适用）

癌症证据不包括诊断日期

7 = 仅医学死亡证明书（DCO） [见说明]

8 = 只有尸检——仅在尸检时发现恶性肿瘤 [见说明]

无相关信息

9 = 不详

说明：

单独凭借死亡证明书（编码 7）或者尸检（编码 8）登记的癌症通常包含在登记当年的癌症发病率统计中。

DCO 或者尸检发现的病例真实诊断日期和生存期不详，因此不能被包含在生存分析中，少数癌症登记处甚至不记录 DCOs 或者尸检发现的癌症。

然而，如果你们登记处已经登记了数据提交时间范围内的这些病例，那么**请务必将DCO和尸检发现的病例包含在本研究要提交的数据内**，使得质量控制时具有可比性。

变量 20 ICD-O-3 解剖位置

四位字符变量 字母数值型

肿瘤部位（解剖位置）应当根据国际疾病分类标准肿瘤专辑第三版（ICD-O-3¹）编码。

请提供完整的 4 位 ICD-O-3 编码，但是请不要带小数点（“.”）。例如，肝癌将是 C220 或者 C221，前列腺癌将是 C619。经过这个细微改动以后，检索癌症的解剖学部位应被编码为：

胃癌： C160-C166; C168-C169

结肠癌： C180-C189; C199

直肠癌： C209; C210-C212, C218

注意：包括肛门和肛管，C210-C218

肝癌： C220-C221

注意：包括肝内胆管，C221.

肺癌 C340-C343; C348-C349

注意：本研究中不包含气管（ICD-O-3 C339）

乳腺癌： C500-C506; C508-C509

宫颈癌： C530-C531; C538-C539

卵巢癌： C480-C482; C569; C570-C574; C577-C579

注意：包含来自腹膜和腹膜后腔的(C480-C482) 来源于输卵管的高级别浆液型的恶性肿瘤(C570)。

注意：包括其它和未特指的女性生殖气管，C577-C579

前列腺癌： C619

白血病： **白血病应当在形态学编码（变量 21）的基础上选择，而不要在解剖位置的基础上作选择。任何有效 ICD-O-3 位置学编码都是可用的，如上所述，请不要带小数点。**

变量 21 ICD-O-3 形态学

四位数字变量，数值型

肿瘤形态学应当根据国际疾病分类肿瘤学专辑（ICD-O-3）编码。

对于显微镜下证实的肿瘤，ICD-O-3 形态学编码范围为：

8000-9989

对于血液学恶性肿瘤，ICD-O-3 形态学编码范围为：

9590-9989

这个形态学编码范围应当被用来选择所有白血病和其它血液学恶性肿瘤包括淋巴瘤。儿童和成年人的编码范围一致。

白血病是**唯一**需要使用ICD-O-3 形态学编码来选择病例提交数据的检索癌症：所有ICD-O-3 形态学编码在 9590-9989 范围内的肿瘤记录都应被包含在内，**其它 9 种检索肿瘤**都应当在ICD-O-3位置学编码（变量 20）的基础上选择。

成年人白血病的选择和分组将基于血液病学家、病理学家和流行病学家在 HAEMACARE 工作组⁶⁻⁸上一致达成的分类。用来编码和报告血液学肿瘤的 HAEMACARE 指南有英语⁹和西班牙语版本（如有需要请联系我们）。

儿童白血病将与成年人白血病的分组不同，是基于在国际儿童癌症分类（ICCC-3）第三版¹⁰。ICCC-3 分组是基于 ICD-O-3 形态学编码。

对于**实体肿瘤未经显微镜下证实**，您应当使用：

9999 在本研究中代表数据缺失，不是 ICD-O-3 编码

注意：这个编码只对实体肿瘤有效，不包括白血病。**白血病根据形态学定义**，因此它们必须有一个在 9590-9989 范围内形态学编码。

变量 22 动态

一位数字变量 数值型

生存分析将仅包含浸润性的、原发的恶性肿瘤。然而，我们也将报告每种癌症肿瘤行为学的分布。这将使得区域间诊断活动强度的比较成为可能，例如登记的宫颈癌女性中原位癌的比例。

因此，假如你们登记处收集良性（动态编码为 0）、动态不确定（1）或者原位癌（2），例如宫颈的原位癌，提交数据时请将所有在检索部位诊断的肿瘤（动态编码 0-3）记录包含在内。

请不要因为肿瘤动态而入选或排除肿瘤记录。

肿瘤动态应当按照国际肿瘤疾病分类（ICD-O-3）第三版编码。肿瘤动态的编码在所有 ICD-O 的修正版中是一样的。

编码 意义

- 0 = 良性
- 1 = 不确定良性恶性
- 2 = 原位癌
- 3 = 恶性，原发灶

以下的编码包含在 ICD-O-3 中，但是癌症登记处不经常使用。为了完整起见，列在下面。

假如您的数据的确包含编码 6 和 9，请不要在数据提交前重新编码，请提供给我们一份说明，说明这些编码在您的数据中是如何使用的。

编码 意义

- 6 = 恶性，继发部位
- 9 = 恶性，不确定原发还是继发部位

提供诊断时的分期数据（变量 23-27）是可选的。

“肿瘤分期”描述诊断时癌症已经扩散到什么程度。它是患者生存（预后因素）的决定性因素。癌症患者诊断时分期的知识在国际生存比较中越来越重要¹¹⁻¹⁵。对于可能的地区，我们将进行与诊断时分期有关的生存分析。

在收集了分期数据的登记处中，至少在近些年，十种检索癌症的质量满意度信息大部分是可用的。

假如你们登记处没有收集关于肿瘤分期的数据，您应该将每条肿瘤记录的所有分期变量（变量 23-27）赋值为 9, 99 或 999（见 45-46 页清单）。您不需要再阅读关于这些变量的其他细节。但请您确保您的数据符合总的说明，这在 45-46 页中有总结。

假如您的确提交了分期数据，请仅提供 2001 年 1 月 1 日以后诊断的患者分期数据。对于 2001 年以前诊断的患者，请在每条记录中将这些变量（变量 23-27）赋值为 9, 99 或 999，具体填写取决于这些变量（见 45-46 页清单）设置了多少位数。2001 年之前确诊的患者分期数据将不包含在质量控制或生存分析内。

世界各地有许多不同的编码方案用来在登记处中对肿瘤分期进行归类。¹⁶

我们将尝试获取基于至少一种世界通用的分期分类标准数据，使得按诊断时分期的生存分析成为可能。

SEER Summary Stage 2000

FIGO stage（宫颈和子宫）

TNM（病理学和临床）

肿瘤大小

Condensed TNM

检查的淋巴结数量

Duke's 分期（结直肠）

因肿瘤良性的淋巴结数量

细节在以下页面中提供。

如果您希望提交诊断时的分期数据，但您的分期数据既不是按照TNM第七版编码又不是按照上表所示的其他分类标准编码，请在提交数据前联系我们。

一位数字变量，数值型

SEER摘要分期 2000 是对分期的一个简要分类，由美国国家癌症研究所的监测、流行病学和最终结果（SEER）项目开发。北美中央癌症登记协会使用SEER摘要分期 2000¹⁷

（<http://seer.cancer.gov/tools/ssm/>）。这里我们借用一下其文字介绍。此分期自 2001 年 1 月 1 日起就在美国和加拿大得到使用。

“摘要分期”是最基本根据癌症扩散程度分类的方法。摘要分期使用的医疗记录提供的所有信息，它是一个将最精确的临床证据和病理证据结合起来对疾病程度的反映。许多基于人群的癌症登记处会报告登记病例的摘要分期。由于分期的类别跨度比较宽，可以用来衡量癌症控制的进展。

我们认为北美的登记处将直接提供 2001-2003 年诊断病例的SEER 摘要分期编码，但是 2004 年以后诊断的病例则提供协作分期（<http://seer.cancer.gov/tools/collabstaging>）。如果你们登记处将以一种不同的方式提供的SEER摘要分期 2000，请告知我们您所使用的程序。这些分期方案的跨时间的可比性在以下的SEER网页中有说

明：http://seer.cancer.gov/seerstat/variables/seer/yr1973_2009/lrd_stage/index.html

根据癌症的扩散方式，癌症的区域扩散被分为几类：

编码 意义

- 0 = 原位癌
- 1 = 限于局部
- 2 = 仅通过直接途径局部蔓延
- 3 = 仅涉及局部淋巴结
- 4 = 直接途径和淋巴结局部扩散
- 5 = 局部性，NOS（未特指）- 癌症局部扩散，扩散途径不详
- 7 = 转移到远端部位或远端淋巴结
- 9 = 扩散或转移不详或这是个仅医学死亡证明书或仅尸检发现的案例

TNM 分期（变量 24-29）

可选

肿瘤-淋巴结-转移（TNM）诊断时分期，像 SEER Summary Stage 2000 一样，结合使用临床和病例证据来进行分期。

TNM分期法由国际癌症控制联盟（UICC）出版。我们将使用TNM手册第7版¹⁸。这跟美国癌症联合委员会（AJCC）2009年¹⁹出版的分类方法是一致的。如果您有通过早期TNM版本编码的分期数据，请在提交数据前联系我们。

TNM分期的三个组成部分是：肿瘤大小（T）；局部淋巴结状态，例如被肿瘤浸润的淋巴结的延伸（N），是否有远端转移（疾病由原发灶扩散到远端器官）（M）。

TNM分期数据可基于病理学依据（“p”）或临床依据（“c”）。

变量 24 病理学 T

可选

一位数字变量，数值型

此变量根据肿瘤的物理大小编码。

对于 9 种实体检索癌症（即除去白血病），每种类型的分期 pT1、pT2、pT3、pT4 最多存在 4 种亚分类（a, b, c, d）。这些亚分类应与其父分类的编码方式一致，例如：pT1a 应当被编码为“1”，与 pT1 一致。

仅对于宫颈癌，还存在附加亚分类：pT1a1, pT1a2, pT1b1 和 pT1b2 应当如 pT1 一样被编码为“1”。同理，亚分类 pT2a1 和 pT2a2 应当如 pT2 一样被编码为“2”。

将使用以下编码：

编码 意义

0 = pT0 - 无原发肿瘤组织学证据

1 = pT1 - 肿瘤各维度小于等于 2cm

此编码也应用于亚分类 pT1a, pT1b 和 pT1c 以及亚分类 pT1a1, pT1a2, pT1b1 和 pT1b2（仅宫颈癌）

2 = pT2 - 肿瘤最大维度大于 2cm，但小于等于 5cm

此编码也应用于亚分类 pT2a, pT2b 和 pT2c 以及亚分类 pT2a1, pT2a2（仅宫颈癌）

3 = pT3 - 肿瘤最大维度大于 5cm

此编码也应用于亚分类 pT3a 和 pT3b

4 = pT4 - 无论肿瘤大小，临近器官有转移

此编码也应用于亚分类 pT4a, pT4b, pT4c 和 pT4d

8 = 原位癌

9 = pTX - 未知：原发肿瘤无法组织学评估

变量 25 病理学 N

可选

一位数字变量，数值型

此变量根据淋巴结被肿瘤浸润的程度编码。

对胃，结肠，直肠和乳腺的癌症， pN1, pN2 和 pN3 还存在一些亚分类 (a, b, c)。这些亚分类编码方式与上一级别分类相同，例如： pN2b 应该编码 “2”，与 pN2 编码相同。

将使用以下编码：

编码 意义

0 = pN0 - 组织学上无区域淋巴结受累

1 = pN1 - 不同癌症分类内容不同（见 TNM 指导手册¹⁸）

对于结肠，直肠以及乳腺，此编码也应有于亚分类 pN1a, pN1b 和 pN1c

2 = pN2 - 不同癌症分类内容不同（见 TNM 指导手册¹⁸）

对于结肠，直肠以及乳腺，此编码也应有于亚分类 pN2a, pN2b 和 pN2c

3 = pN3 - 不同癌症分类内容不同（见 TNM 指导手册¹⁸）

对于胃癌，此编码也应有于亚分类 pN3a 和 pN3b

对于乳腺癌，此编码也应有于亚分类 pN3a, pN3b 和 pN3c

9 = pNX - 未知：区域淋巴结受累无法组织学评估

变量 26 病理学 M

可选

一位数字变量，数值型

此变量编码肿瘤是否有远处转移的信息。

对结肠，直肠和前列腺的癌症， pM1 还存在一些亚分类 (a, b, c)。应该编码为“1”，与 pM1 编码相同。

将使用以下编码：

编码 意义

1 = pM0 - 有远端转移，已显微镜下确认

此编码同样应用于亚分类 pM1a, pM1b 和 pM1c

9 = 不详 - 这不是 TNM 编码（如下）。请当病理 M 状态不明时使用编码 9。

注释：

编码 pM0 和 pMx 在第七版 TNM¹⁸ 中不适用，见下文变量 29 的说明。

变量 27-29 TNM 临床分期

可选

肿瘤各期的临床数据可以来自于临床检查或肿瘤的影像（X 线，CT，核磁等）。

如果没有经过手术或有创性诊断程序，临床数据可能是唯一是能提供肿瘤分期的数据。

这些变量是可选的。如果您选择不提供它们，请在每条记录中将它们赋值为 9。

变量 27 临床 T

可选

1 位数字变量，数值型

此变量提供肿瘤大小方面的信息。

仅在没有病理数据（变量 24）的情况下提供肿瘤 T 期的临床数据。

对于 9 种实体检索癌症（即除去白血病），肿瘤分期中 cT1, cT2, cT3 和 cT4 最多分别有 4 种亚分类（a, b, c, d）。这些亚分类编码方式应该与上一级别分类相同，例如：cT1a 应该编码为“1”，与 cT1 编码相同。

只有宫颈癌有其他亚分类：cT1a1, cT1a2, cT1b1 和 cT1b2 应该记录为“1”，而 cT2a1 与 cT2a2 应该编码为“2”。

将使用以下编码：

编码 意义

0 = cT0 - 无原发肿瘤证据

1 = cT1 - 肿瘤各维度小于等于 2cm

此编码也应用于亚分类 cT1a, cT1b 和 cT1c 以及亚分类 cT1a1, cT1a2, cT1b1 和 cT1b2（仅宫颈癌）

2 = cT2 - 肿瘤最大维度大于 2cm，但小于等于 5cm

此编码也应用于亚分类 cT2a, cT2b 和 cT2c 以及亚分类 cT2a1, cT2a2（仅宫颈癌）

3 = cT3 - 肿瘤最大维度大于 5cm

此编码也应用于亚分类 cT3a 和 cT3b

4 = cT4 - 无论肿瘤大小，临近器官有转移

此编码也应用于亚分类 cT4a, cT4b, cT4c 和 cT4d

8 = 原位癌

9 = cTX - 未知：原发肿瘤无法评估

变量 28 临床 N

可选

1 位数字变量，数值型

此变量根据淋巴结被肿瘤浸润的程度编码。

仅在没有病理数据（变量 25）情况下提供肿瘤 N 期的临床数据。

对胃，结肠，直肠和乳腺的癌症，还有一些 cN1，cN2 和 cN3 的亚分类（a，b，c）。这些亚分类编码方式应该与上一级别分类相同，例如：cN2b 应该编码“2”，与 cN2 编码相同。

将使用以下编码：

编码 意义

0 = cN0 - 临床上肿瘤无区域淋巴结受累

1 = cN1 - 不同肿瘤分类不同（见 TNM 指导手册¹⁸）

对于结肠，直肠以及乳腺，此编码也应有于亚分类 cN1a，cN1b 和 cN1c

2 = cN2 - 不同肿瘤分类不同（见 TNM 指导手册¹⁸）

对于结肠，直肠以及乳腺，此编码也应有于亚分类 cN2a，cN2b 和 cN2c

3 = cN3 - 不同肿瘤分类不同（见 TNM 指导手册¹⁸）

对于胃癌，此编码也应有于亚分类 cN3a 和 cN3b

对于乳腺癌，此编码也应有于亚分类 cN3a，cN3b 和 cN3c

9 = cNX - 未知：区域淋巴结受累无法评估

1 位数字变量，数值型

此变量提供肿瘤是否有远处转移到信息。

对结肠，直肠和前列腺的癌症，存在一些 cM1 的亚分类（a, b, c），应编码为“1”，与 cM1 编码相同。

将使用以下编码：

编码 意义

0 = cM0 - 远处无转移

1 = cM1 - 远处转移。

注释：

如果临床医生没有记录是否有远处转移，根据 TNM 第 7 版的说明，则假定没有远处转移(cM0)：应编码为“0”。

编码“MX”在 TNM 手册之前的版本中表示远处转移未知。但是远处转移的临床评估应该仅基于体格检查，所以 cMX 不再认为是合适的编码。

在 TNM 手册第 7 版¹⁸中编码 cMX 无效。

变量 30-32 缩略 TNM 分期

可选

欧洲癌症登记处网络²⁰针对肿瘤编码制定了缩略版 TNM 分期，其中 T，和/或 N，和/或 M 不被明确赋值。缩略版 TNM 分期是基于 TNM 分期手册第 6 版²¹。

缩略版 TNM 分期的数据仅在**既没有病理 TNM 数据（变量 24-26）也没有临床 TNM 数据（变量 27-29）**情况下提供。

缩略版 TNM 分期与经常用于癌症登记的简化分期相一致，其中疾病的程度被分为局部的、区域性的和扩散的。

变量 30 缩略 T

可选

1 位数变量，数值型

编码 意义

1 = L - 局部的疾病

局部的疾病指胃，结肠，直肠，肝，肺，宫颈及前列腺的癌症的 T1 和 T2 期；乳腺癌的 T1，T2 和 T3 期；卵巢癌的 T1 期（注：卵巢癌 T2 期被认为是激进性的）

2 = A - 激进性疾病

激进性疾病指胃，结肠，直肠，肝，肺，宫颈及前列腺的癌症的 T3 和 T4 期；乳腺癌的 T4 期（注：乳腺癌的 T3 期被认为是局部的的）；卵巢癌的 T2 和 T3 期

9 = X - 无法评估：没有肿瘤大小分类的信息

变量 31 缩略 N

可选

1 位数变量，数值型

编码 意义

1 = N0 - 肿瘤无区域淋巴结受累

2 = N+ - 区域淋巴结受累

9 = NX - 无法评估：无淋巴结状态信息

变量 32 缩略 M

可选

1 位数变量，数字型

此编码基于最佳可得信息：临床，仪器检测或病理信息。即便没有远处转移的病理确定信息，临床标志或结果是判断肿瘤已有远处转移（M+）分期的充足证据。

编码 意义

1 = M0 - 无远处转移

2 = M+ - 远处转移

9 = MX - 无法评估：无法得知是否有远处转移

1 位数变量，数值型

此变量为可选项。如果没有信息提供，请把每条记录中的该变量编码为“9”。

杜克分期²²是仅针对结肠和直肠癌症特殊分期方法。其他所有的检索癌症请编码为“9”

因为 TNM 分期更详细，所以优先使用 TNM 分期。

杜克分期应仅在没有 TNM 分期数据（变量 24-29）情况下报告。

杜克分期调整后²³加入远处转移的分类（D），以及直接转移的 B 和 C 的亚分类。调整后的杜克分期不再建议在临床上应用，但仍然被广泛应用。

编码 意义

1 = Dukes' A 期（相当于 T1N0M0 或 T2N0M0）

2 = Dukes' B 期（相当于 T3N0M0 或 T4N0M0）

3 = Dukes' C 期（相当于 N1M0 的任一 T 期或 N2M0 的任一 T 期）

4 = Dukes' D 期（相当于 M1 的任一 T 期和任一 N 期）

9 = Dukes' 分期不详：无杜克分期的信息

变量 34 FIGO 分期

可选

1 位数字变量，数值型

此变量为可选项。如果没有信息提供，请把每条记录中的该变量编码为“9”。

FIGO 分期²⁴是针对宫颈，卵巢以及其他生殖系统癌症的特殊分期方法。**其他所有的癌症索引请编码为“9”**

因为 TNM 分期更详细，所以优先使用 TNM 分期。

FIGO 分期**仅**在没有 TNM 分期数据（变量 24-29）情况下报告。

FIGO 分期提供 5 大类：

- 0 期 原位癌（多见于宫颈癌）
- I 期 癌症局限于原发器官
- II 期 癌症浸润临近器官或组织
- III 期 癌症在骨盆内远处结节或组织转移
- IV 期 远处转移

宫颈癌（C530, C531, C538, C539）

编码 意义

0 = FIGO 0 期 - 原位癌

1 = FIGO I 期 - 癌症局限于子宫颈（忽略扩散到子宫内膜的肿瘤）

此编码也应用于 I A, I A1, I A2, I B, I B1, I B2 的亚分类

2 = FIGO II 期 - 肿瘤浸润子宫外但未到盆腔壁或阴道下部三分之一

此编码也应用于 II A, II A1, II A2, II B 的亚分类

3 = FIGO III 期 - 肿瘤延伸至盆腔壁或阴道下部三分之一，或引起肾盂积水

此编码也应用于 IIIA, III B 的亚分类

4 = FIGO IVA 期 - 肿瘤浸润至膀胱粘膜或直肠，或延伸超出盆腔

5 = FIGO IVB 期 - 远处转移

6 = FIGO 分期不详

卵巢癌（C569）或输卵管癌（C570）

编码 意义

1 = FIGO I 期 - 肿瘤局限于单侧或双侧卵巢*

此编码也应用于 I A, I B, I C 的亚分类

2 = FIGO II 期 - 肿瘤浸润单侧或双侧卵巢*, 同时延伸至盆腔

此编码也应用于 II A, II B, II C 的亚分类

3 = FIGO III 期 - 肿瘤浸润单侧或双侧卵巢*, 同时显微镜下确认盆腔外和/或区域淋巴结有腹膜转移。

此编码也应用于 III A, III B, III C 的亚分类

4 = FIGO IV 期 - 腹膜腔外的远处转移（除外腹膜转移）

9 = FIGO 分期不详

*对于输卵管癌（C570），请用“输卵管”替换“卵巢”。

变量 35 肿瘤大小

可选

1 至 3 位数字变量，数值型

报告肿瘤大小（最大直径）时必须以**毫米**作为单位，报告整数。

对于**乳腺癌**，如果可行的话，肿瘤大小应该基于组织学检测。

对于**肺癌**，肿瘤大小可由影像学检测报告。

肿瘤最大尺寸的大小很难描述。范围在 1 - 300mm (1mm – 30cm) 的数值均可接受。例如，若一个肿瘤的最大直径是 35mm (3.5cm)，应该编码为“35”。

编码 意义

1-300 = 肿瘤最大直径，**单位毫米**

999 = 肿瘤最大直径不详，或最大直径不适用（**白血病**），或变量未提供

“0”不是有效的肿瘤大小记录。如数据缺失，此变量记录为 999。

变量 36 淋巴结检测数目

可选

1 至 2 位数字变量，数值型

记录**淋巴结检测**的具体数目，如病理记录所记录：有效范围 0-98。

以下情况下，请将**每条记录**中为此变量赋值为 99：

- 您选择不提供此变量，或
- 在白血病的数据文件中

对于九个**实体肿瘤**，以下情况下此变量应编码为 99：

- 无可用的信息，或
- 未作病理检查

变量 37 淋巴结受累数目

可选

1 至 2 位数字变量，数字型

记录**受累淋巴结(含有肿瘤细胞的淋巴结)**的精确数目，如病理记录所记录的：有效值范围 0-98。

以下情况下，请将**每条记录**中为此变量赋值为 99：

- 您选择不提供此变量，或
- 在白血病的数据文件中

对于九个**实体肿瘤**，以下情况下此变量应编码为 99：

- 无可用的信息，或
- 淋巴结受累但是受累淋巴结**数量**不详，或
- 未作病理检查

4 变量的有效数值

变量名	描述	数字或字符个数	有效值	缺失值记录
VAR1	国家	4	见第 6 页中的列表	不允许
VAR2	登记处	1, 2 或 3	请用提供的数值*	不允许
VAR3	个人编码	最大值为 15	数字或字母	不允许
VAR4	肿瘤编码	最大值为 10	数字或字母	不允许
VAR5	IARC检查标志 ¹	1	1, 2, 3, 4	9
VAR6	性别	1	1, 2	9
VAR7	地区 ²	最大值为 5	数值型 (见第 12 页)	99999
VAR8	民族 ²	1 或 2	数值型 (见第 13 页)	99
VAR9	出生日期-日	1 或 2	1-31	99
VAR10	出生月份	1 或 2	1-12	99
VAR11	出生年份	4	1895-最近年份	9999
VAR12	诊断日期-日	1 或 2	1-31	99
VAR13	诊断月份	1 或 2	1-12	99
VAR14	诊断年份	4	1995-最近年份	9999
VAR15	末次得知生命状态	1	1, 2, 3	9
VAR16	末次得知生命状态日期-日	1 或 2	1-31	99
VAR17	末次得知生命状态月份	1 或 2	1-12	99
VAR18	末次得知生命状态年份	4	1995-最近年份	9999
VAR19	诊断依据	1	1, 2, 3, 4, 5, 6, 7, 8	9
VAR20	ICD-O-3 解剖学部位	4	数字或字母	不允许
VAR21	ICD-O-3 形态学	4	8000-9989	9999
VAR22	动态	1	0, 1, 2, 3, 6, 9	不允许
VAR23	SEER Summary Stage 2000	1	0, 1, 2, 3, 4, 5, 7	9
VAR24	病理 T	1	0, 1, 2, 3, 4, 8	9
VAR25	病理 N	1	0, 1, 2, 3	9
VAR26	病理 M	1	1	9

VAR27	临床 T	1	0, 1, 2, 3, 4, 8	9
VAR28	临床 N	1	0, 1, 2, 3	9
VAR29	临床 M	1	0, 1	9
VAR30	缩略 T	1	1, 2	9
VAR31	缩略 N	1	1, 2	9
VAR32	缩略 M	1	1, 2	9
VAR33	Dukes' 分期 ³	1	1, 2, 3, 4	9
VAR34	FIGO 分期 ⁴	1	0, 1, 2, 3, 4, 5	9
VAR35	肿瘤大小	1, 2 或 3	1-300	999
VAR36	淋巴结检查数目	1 或 2	0-98	99
VAR37	淋巴结受累数目	1 或 2	0-98	99

* 此编码将会单独提供给各参与登记处

1 可选 - 如果登记处不提供此变量，请将每条记录中此变量编码为“9”（见最后一列）。

2 可选 - 如果登记处提供此变量，请提前与我们确认有效的代码；否则此处编码为 99999（变量 7）或 99（变量 8）（见最后一列）

3 仅对于结直肠肿瘤，其他肿瘤请填写 9（见最后一列）

4 仅对于宫颈和卵巢肿瘤，其他肿瘤请填写 9（见最后一列）

5 参考文献

1. Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin LH, Parkin DM, Whelan SL, eds. *International Classification of Diseases for Oncology (ICD-O), third edition*. Geneva: World Health Organisation; 2000
2. World Health Organisation. *International Classification of Diseases, 1975, 9th revision*. Geneva: WHO; 1977
3. World Health Organisation. *International statistical classification of diseases and related health problems. Tenth revision*. Geneva: WHO; 1994
4. Percy C, Van Holten V, Muir CS, eds. *International Classification of Diseases for Oncology (ICD-O), second edition*. Geneva: World Health Organisation; 1990
5. Woods LM, Rachet B, Ellis L, Coleman MP. Full dates (day, month, year) should be used in population-based cancer survival studies. *Int J Cancer* 2012; **131**: E1120-E1124
6. Sant M, Allemani C, Tereanu C, De Angelis R, Capocaccia R, Visser O, Marcos-Gragera R, Maynadié M, Simonetti A, Lutz J-M, Berrino F. Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood* 2010; **116**: 3724-34
7. Marcos-Gragera R, Allemani C, Tereanu C, De Angelis R, Capocaccia R, Maynadié M, Luminari S, Ferretti S, Johannesen TB, Sankila R, Karjalainen-Lindsberg M-L, Simonetti A, Martos MC, Raphael M, Giraldo P, Sant M. Survival of European patients diagnosed with lymphoid neoplasms in 2000-2002: results of the HAEMACARE project. *Haematologica* 2011; **96**: 720-8
8. Maynadié M, De Angelis R, Marcos-Gragera R, Visser O, Allemani C, Tereanu C, Capocaccia R, Giacomini A, Lutz J-M, Martos C, Sankila R, Johannesen TB, Simonetti A, Sant M. Survival of European patients diagnosed with myeloid malignancies: a HAEMACARE study [epub ahead of print]. *Haematologica* 2012
9. Sant M, Karjalainen-Lindsberg M-L, Maynadié M, Raphael M, Ferretti S, Giacomini A, Tereanu C, Giraldo P, Marcos-Gragera R, Martos C, Lutz J-M, Visser O, HAEMACARE Working Group. Manual for coding and reporting haematological malignancies. *Tumori* 2010; **96 (Appendix)**: 1-32
10. Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, third edition. *Cancer* 2005; **103**: 1457-67
11. Sant M, Allemani C, Capocaccia R, Hakulinen T, Aareleid T, Coebergh JWW, Coleman MP, Grosclaude PC, Martinez-Garcia C, Bell CMJ, Williams EMI, Berrino F, EURO CARE Working Group. Stage at diagnosis is a key explanation of differences in breast cancer survival across Europe. *Int J Cancer* 2003; **106**: 416-22
12. Ciccolallo L, Capocaccia R, Coleman MP, Berrino F, Coebergh JWW, Damhuis RAM, Faivre J, Martinez-Garcia C, Møller H, Ponz de Leon M, Launoy G, Raverdy N, Williams EMI, Gatta G. Survival differences between European and US patients with colorectal cancer: role of stage at diagnosis and surgery. *Gut* 2005; **54**: 268-73
13. Sankaranarayanan R, Swaminathan R, Brenner H, Chen K, Chia KS, Chen JG, Law SCK, Ahn YO, Xiang YB, Yeole BB, Shin HR, Shanta V, Woo ZH, Martin N, Sumitsawan Y, Sriplung H, Barboza AO, Eser S, Nene BM, Suwanrungruang K, Jayalekshmi P, Dikshit R, Wabinga H, Esteban DB, Laudico A, Bhurgri Y, Bah E, Al-Hamdan N. Cancer survival in Africa, Asia, and Central America: a population-based study. *Lancet Oncol* 2010; **11**: 165-73
14. Maringe C, Walters S, Butler J, Coleman MP, Hacker N, Hanna L, Mosgaard BJ, Nordin A, Rosen B, Engholm G, Gjerstorff ML, Hatcher J, Johannesen TB, McGahan CE, Meehan D, Middleton R, Tracey E, Turner D, Richards MA, Rachet B, ICBP Module 1 Working Group. Stage at diagnosis and ovarian cancer survival: evidence from the International Cancer Benchmarking Partnership. *Gynec Oncol* 2012; **127**: 75-82
15. Allemani C, Sant M, Weir HK, Richardson LC, Baili P, Storm H, Siesling S, Torrella-Ramos A, Voogd AC, Aareleid T, Ardanaz E, Berrino F, Bielska-Lasota M, Bolick S, Cirilli C, Colonna M, Contiero P, Cress RD, Crocetti E, Fulton JP, Grosclaude P, Hakulinen T, Izarzugaza I, Malmström P, Peignaux K, Primic-Zakelj M, Rachtan J, Safaei Diba C, Sanchez M-J, Schymura MJ, Shen T, Traina A, Tryggvadóttir L, Tumino R, Velten M, Vercelli M, Wolf HJ, Woronoff A-S, Wu X, Coleman MP. Breast cancer survival in the US and Europe: a CONCORD high-resolution study [Epub ahead of print]. *Int J Cancer* 2012
16. Walters S, Maringe C, Butler J, Brierley JD, Rachet B, Coleman MP. Comparability of stage data in cancer registries in six countries: lessons from the International Cancer Benchmarking Partnership [Epub ahead of print]. *International Journal of Cancer* 2012
17. Young JL, Roffers SD, Ries LAG, Fritz AG, Hurlbut AA. *SEER Summary Staging Manual - 2000: Codes and coding instructions*. Bethesda MD, National Cancer Institute, 2001.
18. Sobin LH, Gospodarowicz M, Wittekind C. *TNM Classification of Malignant Tumours*. 7th edn. New York: John Wiley & Sons; 2009
19. Edge SB, Byrd DR, Carducci MA, Compton CC, Fritz AG, Greene F, Trotti A, eds. *American Joint Committee on Cancer - cancer staging manual*. 7th edn. New York, NY: Springer; 2009
20. Berrino F, Brown C, Moller T, Sobin LH, Faivre J. *ENCR recommendations: condensed TNM for coding the extent of disease*. Lyon, European Network of Cancer Registries, 2002.
21. Sobin LH, Wittekind C, eds. *TNM Classification of Malignant Tumours*. 6th edn. New York: John Wiley & Sons; 2002
22. Dukes CE. The classification of cancer of the rectum. *J Path Bacteriol* 1932; **35**: 323-32
23. Astler VB, Coller FA. The prognostic significance of direct extension of carcinoma of the colon and rectum. *Ann Surg* 1954; **139**: 846-52
24. Benedet JL, Bender H, Jones H, Ngan HY, Pecorelli S. FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. *Int J Gyn Obs* 2000; **70**: 209-62