

关联规则在 14 种食物不耐受健康体检中的应用

钟华 刘启贵 张浩 孟宪云 赵鹤 台文霞

【摘要】目的 以数据挖掘方法中的关联规则法,探索健康人群 14 种食物不耐受的食物之间的关联规律,从而为健康人群提供更为合理的饮食搭配。**方法** 2008 年 1 月至 2010 年 6 月在中国医科大学附属第一医院体检中心进行食物不耐受成年体检者的 2962 例血清标本。采用酶联免疫方法(ELISA)检测人血清中 14 种食物过敏原特异性 IgG 抗体。使用 SPSS16.0 对体检者的食物进行频数分析。使用 SPSS Clementine 12.0 数据挖掘软件,将 14 种食物按照营养学对于食物来源为分类原则并与性别建立关联规则。**结果** 所调查人群中的食物不耐受的前五位主要为动物性食物:蛋类、螃蟹、虾、奶类和鲑鱼(阳性率分别为 32.5%、26.0%、19.8%、13.8% 和 12.0%)。在总体关联中蛋类→奶类与男性的置信度最高(56.06%)、螃蟹→虾与蛋类提升较高(2.07%)。女性出现单阶强关联,男性多为多阶关联。**结论** 较强的阳性关联出现在动物性食物之间,女性与奶类、蛋类和动物海产类为强关联,男性多为多种食物共同关联。

【关键词】 饮食习惯; 健康管理; 食物不耐受; 关联规则

Association rules in health check-up receivers with food intolerance ZHONG Hua*, LIU Qi-gui, ZHANG Hao, MENG Xian-yun, ZHAO Ge, TAI Wen-xia. * Department of Medical Statistics, Dalian Medical University, Dalian 116044, China
Corresponding author: LIU Qi-gui, Email: liuqiguidl@163.com

【Abstract】Objective To explore the relationship among 14 kinds of intolerant food by using association rules to provide evidence for more reasonable diet intake. **Methods** A total of 2962 adults who received medical examination at the First Affiliated Hospital of China Medical University during January 2008 and June 2010 were enrolled in this investigation. ELISA was used to detect allergen-specific IgG antibody of 14 kinds of food. An association rules model was established by SPSS Clementine12.0. In order to find positive intolerance association rules, food frequency was recorded with SPSS 16.0. **Results** Five most commonly seen food intolerance were eggs, crabs, shrimp, milk, and cod (positive rates were 32.5%, 26.0%, 19.8%, 13.8% and 12.0%, respectively). Positive association rules were found in crab and shrimp, or milk and eggs. Single order association was found in female, while multi-order association was identified in male. **Conclusions** Our data show a strong positive association among animal foods. Female may be affected by milk, eggs and seafood, while male may be affected by a variety of foods.

【Key words】 Food intake habits; Health management; Food intolerance; Association rules

食物不耐受是机体免疫系统对特定食物或食物成分不能很好的消化吸收,从而产生食物特异性 IgG 抗体, IgG 抗体与食物颗粒形成免疫复合物,可能引起所有组织发生炎症反应,并表现为全身各系统的症状与疾病^[1]。关联规则是帮助发现大量数

据库中相集之间的关联关系^[2]。本研究欲通过使用 Clementine12.0 数据挖掘软件,建立关联规则挖掘,对 14 种食物不耐受情况的大样本数据进行深入挖掘,进而了解 2962 例健康体检者食物不耐受的具体情况,并建立关联规则。应用此规则对食物不耐受的体检者给予一定的饮食建议与指导。

对象与方法

1. 研究对象:选取所有 2008 年 1 月至 2010 年 6 月中国医科大学附属第一医院体检中心进行食物不耐受成年体检者的 2962 例血清标本,男性占

DOI: 10.3760/cma.j.issn.1674-0815.2012.03.003

作者单位:116044 大连医科大学卫生统计学教研室(钟华、刘启贵);中国医科大学附属第一医院体检中心(张浩、孟宪云、赵鹤、台文霞)

通信作者:刘启贵,Email:liuqiguidl@163.com

第一届“拜美克杯”获奖论文

51.6%, 年龄为(50.6 ± 12.2)岁, 女性占 48.4%, 年龄为(55.7 ± 12.8)岁。

2. 研究方法:(1)试剂及数据处理原理:食物不耐受检测试剂盒(美国 BIOMERICA 公司)酶联免疫方法(ELISA)检测人血清中 14 种食物过敏原特异性 IgG 抗体。根据吸入度值求得标本 IgG 抗体浓度,该值与食物过敏原特异性 IgG 抗体浓度成正比,单位为 U/ml,针对每种过敏原不同浓度的 IgG 抗体解释如下:检测值小于 50 U/ml,结果判定阴性,分级为 0;检测值大于 50 U/ml,结果判定为敏感。并将结果汇总到 SPSS 中并建立数据库。

(2)食物的营养学分类:日常膳食中的食物有两种来源:来自植物的食物,如谷类、豆类、硬果类、蔬菜水果类、植物油等;来自动物的食物,如肉类、脏腑类、动物海产类、禽类、蛋类、乳类、动物油脂等^[3]。将 14 种食物根据以上原则分为:植物性食物包含粮食类食物、豆类及坚果类食物、蔬菜、水果类食物、食用菌类;动物性食物包含肉类、脏腑类、动物海产类、蛋类食物、乳类食物。

(3)关联的建立过程:根据食物营养学中食物来源的分类原则,分别对不同种类的食物建立关联规则。以关联规则中的支持度、置信度和提升比^[4-5]这三个方面探讨并分析这 2962 例健康体检者对于食物不耐受种类和性别特征的具体分布情况。建立过程见图 1。

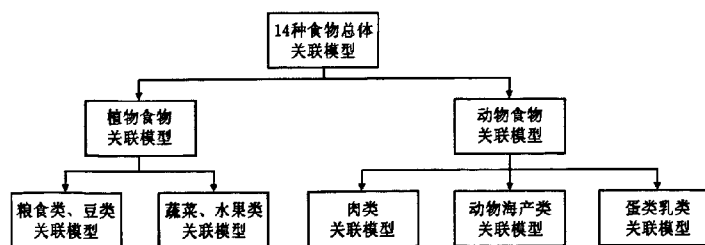


图 1 关联规则依照营养学分类的流程图

(4)统计学方法:数据处理分别使用 SPSS16.0 和 Clementine12.0 进行分析。使用 SPSS16.0 对 14 种食物进行频数分析,观察其阳性频数。使用 Clementine 12.0,对 14 种食物和性别进行关联规则研究。建立了根据 Apriori 算法的关联规则。关联规则设定置信度大于 40%,支持度大于 5%,不够条件者剔除。分析结果着重讨论单阶和二阶关联规则。

(5)先验(Apriori)算法:Apriori 算法 1993 年由 R. Agrawal 和 Srikant 提出来的 Apriori 算法是关联规则挖掘的一个经典算法^[4-5]。这是一个基于两阶

段频数集思想的方法,将关联规则挖掘算法的设计可以分解为两个子问题:找出数据集中所有频繁项目集合和根据频繁项目集合和最小置信度阈值产生出关联规则模型^[6]。

结 果

1. 频数分布情况:在 14 种食物中,受试人群敏感程度由高到低为:蛋类 > 螃蟹 > 虾 > 奶类 > 鳕鱼 > 大豆 > 西红柿 > 玉米 > 牛肉 > 鸡肉 > 大米 > 蘑菇 > 小麦 > 猪肉。其中,蛋类、螃蟹、虾、奶类、鳕鱼位于前五位,明显高于其他食物,结果见表 1。

2. 14 种食物总体关联规则结果:14 种食物总体关联的结果主要分布在动物类食品上。体现在海产品、蛋类和奶类的关联。其中螃蟹和虾,蛋类同奶类的关联程度最高,见图 2。女性对于螃蟹与虾的阳性反应程度和蛋类与虾的过敏程度是基本相同的。多种食物交互二阶关联下,男性和女性的置信度和提升比比较接近。女性在关联规则中出现的次数要多于男性,见表 2。

由于在(表 2)的关联规则,总体为动物类食物,掩盖了一些关联规则的出现。因此需要依照营养学的食物分类原则对这 14 种食物进行分类,然后再次建立关联规则。

3. 按照营养学分类原则下建立的关联规则:
(1)动物类食物总体关联结果:动物类食物分为:肉类、动物海产类、蛋类和奶类,见表 3。可以发现肉类没有出现在关联之中。女性对于动物海产类食物出现了单阶强关联。对动物海产类过敏的人对蛋类和奶类也会有较强的阳性反应。男性在动物海产类→蛋类/奶类和性别关联中的置信度是高于女性的,见表 3。

(2)奶类与蛋类关联结果:女性对于蛋类和奶类食品阳性反应均产生单阶强关联。在奶类与蛋类的关联中,男性则仍为多阶关联,见表 4。

(3)植物类食物总体关联结果:植物类食物分为:粮食类食物、豆类、蔬菜水果和菌类见表 5。可见粮食类食物和豆类的关联程度最强。而其他的关联性较弱特别是没有性别加入其中,见表 5。

讨 论

1. 关联规则中的食物:支持度 > 5%、置信度 > 40%、提升比 > 1% 的关联规则进入讨论。我们发现

表 1 不同种类食物的不耐受的频数分析(例)

食物种类	阴性反应	阳性反应	阳性率(%)
蛋类	2000	962	32.5
螃蟹	2193	769	26.0
虾	2376	586	19.8
奶类	2552	410	13.8
鳕鱼	2607	355	12.0
大豆	2794	168	5.7
西红柿	2864	98	3.3
玉米	2869	93	3.1
牛肉	2918	44	1.5
鸡肉	2918	44	1.5
大米	2919	43	1.5
蘑菇	2922	40	1.4
小麦	2937	25	0.8
猪肉	2951	11	0.4

表 2 14 种食物总体关联结果(%)

规则	支持度	置信度	提升比
蛋类→奶类与男性	6.69	56.06	1.73
蛋类→奶类	13.84	55.37	1.71
蛋类→奶类与女性	7.16	54.72	1.69
螃蟹→虾与蛋类	7.73	53.71	2.07
螃蟹→虾与男性	9.66	49.65	1.91
螃蟹→虾	19.78	49.15	1.89
螃蟹→虾与女性	10.13	48.67	1.88
虾→螃蟹与蛋类	9.22	45.06	2.28
虾→螃蟹与蛋类与女性	5.20	42.86	2.17
蛋类→鳕鱼与女性	5.60	42.77	1.32
蛋类→虾与螃蟹	9.72	42.71	1.32
蛋类→虾与女性	10.13	40.67	1.25

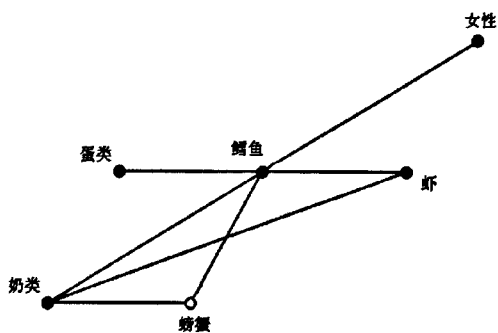


图 2 14 种食物总体关联网络状图

进入关联规则的食物主要分布在动物类食品上,组合主要为:动物海产类→蛋类、蛋类→奶类与动物海产类→奶类,例如:动物海产类→女性,其支持度、置信度均超过 40%,同时提升比也达到了 1.07%。其中典型的是螃蟹→虾的支持度为 19.78%、置信度为 49.15%、提升比为 1.89%。在奶类与蛋类的关联中,从奶类→蛋类和蛋类→奶类两种不同的食物出发的关联规则的支持度是不同的,但其置信度和

万方数据

表 3 动物食物总体关联结果(%)

规则	支持度	置信度	提升比
蛋类→奶类与动物海产类	6.79	61.69	1.92
动物海产类→奶类与蛋类	7.66	54.63	1.30
动物海产类→奶类与男性	6.69	52.02	1.24
动物海产类→奶类	13.84	49.02	1.17
动物海产类→蛋类与男性	14.11	47.85	1.14
动物海产类→奶类与女性	7.16	46.23	1.10
动物海产类→蛋类	32.48	45.84	1.09
动物海产类→女性	48.38	44.87	1.07
动物海产类→蛋类与女性	18.37	44.30	1.05

表 4 奶类与蛋类关联结果(%)

规则	支持度	置信度	提升比
蛋类→奶类与男性	6.69	56.06	1.73
蛋类→奶类	13.84	55.37	1.71
蛋类→奶类与女性	7.16	54.72	1.69
蛋类→女性	48.38	37.96	1.17
奶类→男性与蛋类	14.11	26.56	1.92
奶类→蛋类	32.48	23.60	1.71
奶类→蛋类与女性	18.37	21.32	1.54
奶类→女性	48.38	14.79	1.07

表 5 植物食物总体关联结果(%)

规则	支持度	置信度	提升比
粮食类食物→蘑菇	1.35	25.00	5.04
粮食类食物→西红柿	3.31	16.33	3.29
西红柿→蘑菇	1.35	15.00	4.53
大豆→粮食类食物	4.96	14.29	2.52

提升比接近,说明这两种食物,对于食物不耐受敏感者的影响一致性。由于这组关联规则的较高的支持度和置信度,我们可以用来指导那些有食物不耐受阳性反应的体检者,注意这两种食物的合理饮食。

植物类食物的关联规则没有一项完全符合支持度 > 5%、置信度 > 40%、提升比 > 1% 的指标标准。尽管植物类食物的关联规则的提升比很高,但是其支持度与置信度较低,说明植物间的关联较弱,较高的提升比,这说明对于有蘑菇敏感体质的不耐受体检者应注意粮食类食物的摄取。

2. 关联规则中的性别:表 2~4 的关联规则可以发现:性别主要出现在动物类食物的关联规则中,女性存在单阶强关联,男性主要存在于多阶规则。例如蛋类→奶类与男性规则的置信度和提升比大于蛋类→奶类与女性。一方面,说明了多层关联规则的并不一定是低层的单层关联规则的置信度和提升比之和^[5]。另一方面,也说明了男性在某些食物不耐受程度是与女性接近的,特别是奶类、蛋类和海产品的交叉关联。

女性在食物不耐受的性别关联出现的单阶关

联,奶类→女性的支持度 48.38%、置信度 14.79%、提升比 1.07%。蛋类→女性的支持度 48.38%、置信度 37.96%、提升比 1.17%。且在多阶规则中女性数量中占大多数,这比较符合女性不耐受通常的研究结果^[7,8]。

3. 关联规则应用于食物不耐受研究的作用:健康管理是指一种对个人或人群的健康危险因素进行全面检测、评估与有效干预的活动过程^[9]。将食物不耐受检查引进健康人群体检是对于健康人群出现的慢性过敏性症状病因的亚临床评估手段^[10]。

尽管饮食控制对食物不耐受治疗有着非常积极的效果,但是医护人员不能给出通用的饮食控制方案,更不能给出多种食物的饮食方案^[11]。在现代社会不同种食品交叉食用的现象越来越频繁,食物成分交叉联合致病的可能性逐渐增加^[12]。为了给不同食物不耐受类型健康人提供饮食控制建设性的方案。我们在这次通过健康体检者食物不耐受检查的数据分析中引入了关联规则。通过关联规则,我们可以向体检者建议合理的饮食规划。通过控制某些食物种类的摄入,已达到减少体内 IgG 抗体产生的目的。从预防医学的角度把握本地区人群食物不耐受发生的特点加以合理规划。作为体检工作者适当引入食物不耐受信息评估体系^[13]与饮食健康管理的个案解读,去提醒这些体检者应该注意的食品种类。对于食物不耐受症状较明显的体检者,可以建议其减少摄入关联规则涉及的食品的种类,在一段时间的观察之后判断其食物不耐受症状是否有足够

的减轻^[14]。从宏观与微观两个方向上提高体检者的健康饮食质量。

参 考 文 献

- [1] 谢志贤,刘倩. 食物不耐受与相关性疾病. 中华内科杂志, 2006, 45:150-151.
- [2] 元昌安. 数据挖掘与 SPSS Clementine 应用宝典. 北京: 电子工业出版社, 2009: 176-220.
- [3] 张爱珍. 医学营养学. 北京: 人民卫生出版社, 2009: 100-111.
- [4] 任颖, 李华伟, 吕红. 遗传算法在关联规则挖掘中的应用. 电脑知识与技术: 学术交流, 2009: 4260-4261.
- [5] 戴臧胜, 匡宏波, 谢邦昌. 数据挖掘中的关联规则. 统计研究, 2002(8): 40-42.
- [6] 蔡伟杰, 张晓辉, 朱建秋, 等. 关联规则挖掘综述. 计算机工程, 2001, 27: 31-34.
- [7] 张浩, 钟华, 孟宪云, 等. 健康体检者 2962 例 14 种食物不耐受分析. 中华健康管理学杂志, 2011, 5: 137-139.
- [8] Devonshire V, Duquette P, Dwosh E, et al. The immune system and hormones: review and relevance to pregnancy and contraception in women with MS. Int MS J, 2003, 10: 44-50.
- [9] 中华医学会健康管理学分会, 中华健康管理学杂志编委会. 健康管理概念与学科体系的中国专家初步共识. 中华健康管理学杂志, 2009, 3: 141-147.
- [10] Wüthrich B. Unproven techniques in allergy diagnosis. J Investig Allergol Clin Immunol, 2005, 15: 86-90.
- [11] 曾强, 王晓钟, 马俐华. 健康体检与健康管理服务模式的创新与实践. 中华健康管理学杂志, 2009, 3: 132-134.
- [12] 郑延松, 赵威. 免疫球蛋白 G 抗体介导的食物不耐受. 中华健康管理学杂志, 2010, 4: 183-185.
- [13] 郭季春, 张凌云, 杜强. 健康体检后的个案解读. 中华健康管理学杂志, 2010, 4: 374-375.
- [14] 武留信. 健康信息评估. 中华健康管理学杂志, 2008, 2: 121-122.

(收稿日期: 2012-01-16)

(本文编辑: 袁庆)

· 消息 ·

《中华健康管理学杂志》被国外期刊数据库收录情况

《中华健康管理学杂志》已被西太平洋地区医学索引(Western Pacific Region Index Medicus, WPRIM)、美国化学文摘(CAS)和美国《乌利希期刊指南》(Ulrichweb)收录,特此通告。

《中华健康管理学杂志》编辑部