به نام خدا
talebi_ghane@yahoo.com

کاربرد روشهای آماری در پژوهش های پزشکی

دکتر الهه طالبی قانع
دکترای تخصصی آمار زیستی
عضو هیات علمی دانشگاه علوم پزشکی همدان

METHODS

Biostatistical methods

The study evaluated the impact of specific HLA alleles with a frequency of more than 5% on the outcome of grades III-IV acute GvHD. Outcomes were compared among patient- and donor-specific HLA allele-positive and -negative groups using multivariable competing risk regression analysis,[18] adjusted for clinical factors and HLA allele matching (Table 1). We included separate variables for HLA-A,-C,-B,-DRB1,-DQB1 and -DPB1 allele mismatches in the GvH direction.

To analyze the effect of patient mismatched HLA-C allele, those pairs matched for 1 HLA-C allele and mismatched for another HLA-C allele were extracted. The risk of each patient mismatched HLA-C allele on grades III-IV acute GvHD was compared with the HLA-C allele match. The influences of the level of expression of the patient mismatched HLA-C allotype were assessed as described previously.[19]

The effects of HLA-C allele mismatch combinations were also evaluated using the pairs matched for 1 HLA-C allele and mismatched for another HLA-C allele, and the risk of each HLA-C mismatch combination of grades III-IV acute GvHD was compared with the HLA-C allele match.

Multivariable competing risk regression analyses[18] were conducted to evaluate the impact on acute GvHD and transplant-related mortality. A Cox's proportional hazards regression model

used for
rom the
nine sta-
used for
test to
f Hazard
portional
gender,
, with a

talebi_ghane@yahoo.com

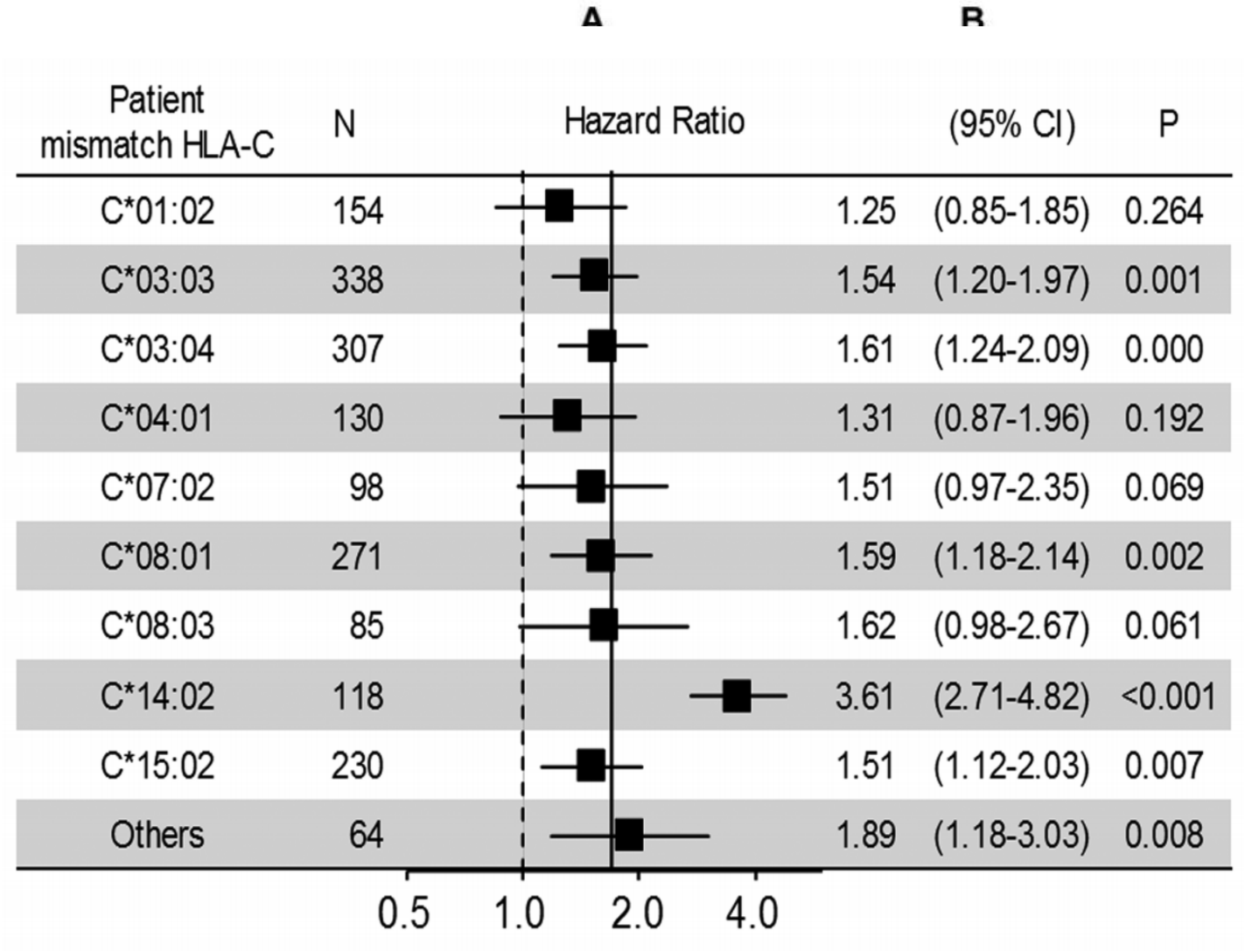


Figure 1. Impact of patient mismatched HLA-C allele on grades III-IV acute GvHD. Results of multivariable competing risk regression analysis for the effect of patient mismatched HLA-C allele on grades III-IV acute GvHD. The hazard ratio (HR) of each mismatched HLA-C allele was estimated by comparison to HLA-C match (n=4825). The solid vertical line at 1.69 indicates the HR of overall HLA-C mismatch in the GvH direction. Others contain patient mismatched HLA-C alleles with fewer than 20 patients, as follows: C*01:03, C*03:14, C*03:23, C*05:01, C*07:04, C*12:02, C*12:03 and HLA-C*14:03.

نقش و مقیاس متغیرها

talebi_ghane@yahoo.com

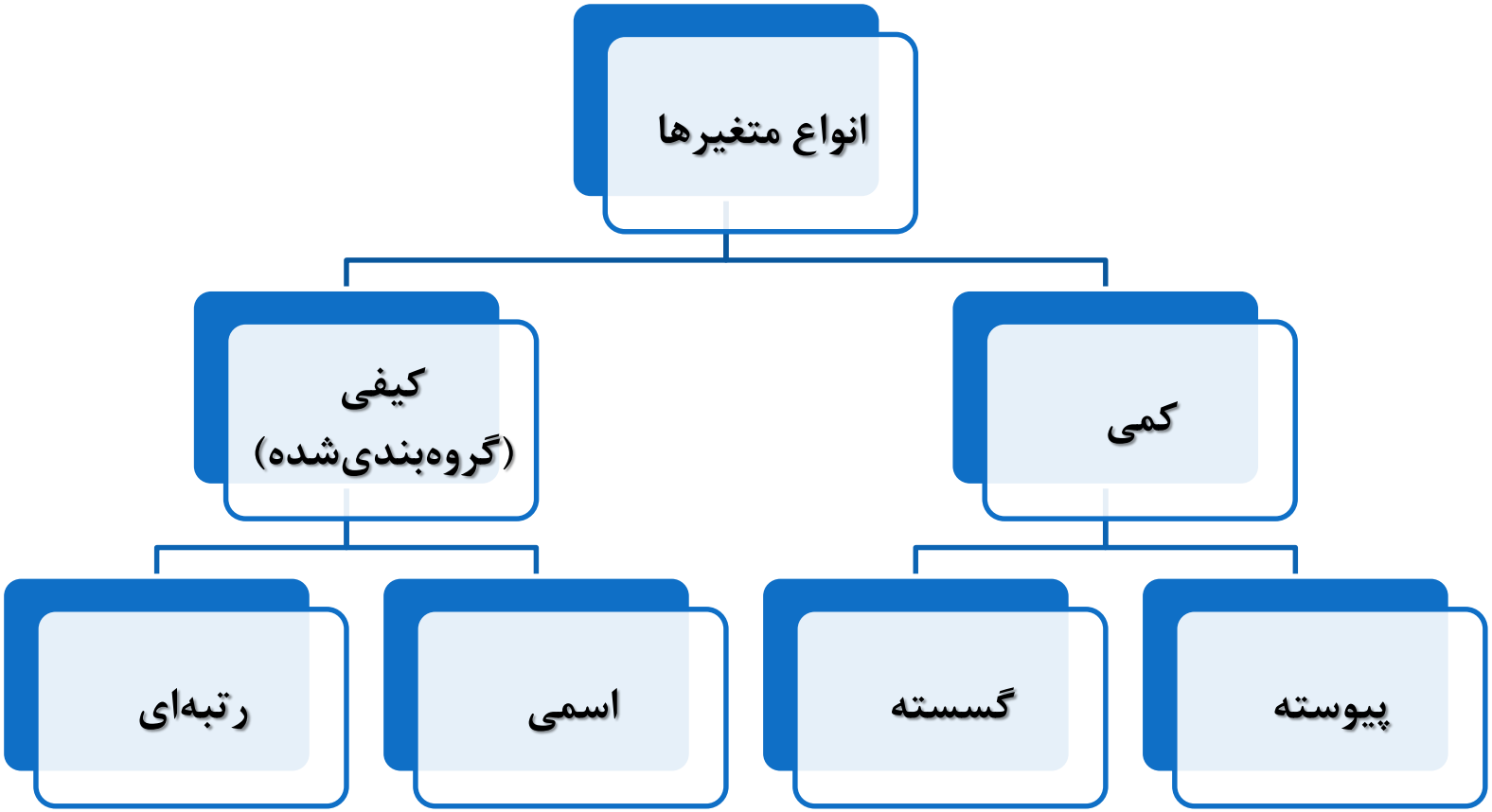
متغیرها

- **متغیرها، عامل یا عواملی هستند که مورد اندازه گیری یا سنجش قرار می‌گیرند و به عبارت دیگر متغیر مشخصه یک فرد، چیز، پدیده مورد نظراست که:**
  - **قابل اندازه گیری بوده**
  - **می تواند مقادیر مختلفی بپذیرد**

لازم است محقق به تعریف و مشخص ساختن آن متغیر برای اندازه گیری و سنجش و تعیین نوع و نقش آن در مطالعه بپردازد.

talebi_ghane@yahoo.com

انواع متغیرها



talebi_ghane@yahoo.com

انواع متغیرها

$CD8^+$ T-cell dose
$< 47.7 \times 10^6$
$> 47.7 \times 10^6$
Total iNKT-cell dose
$> 0.057 \times 10^6$
$< 0.057 \times 10^6$
$CD4^+$ iNKT-cell dose
$> 0.026 \times 10^6$
$< 0.026 \times 10^6$
$CD4^-$ iNKT-cell dose
$> 0.031 \times 10^6$
$< 0.031 \times 10^6$

- مثال:
- اسمی: جنسیت (زن-مرد)
  گروه‌خونی (O-AB-B-...
- رتبه‌ای: شدت یک بیماری (ک...
- گسسته: تعداد دندان‌های پوس...
- پیوسته: قد- وزن

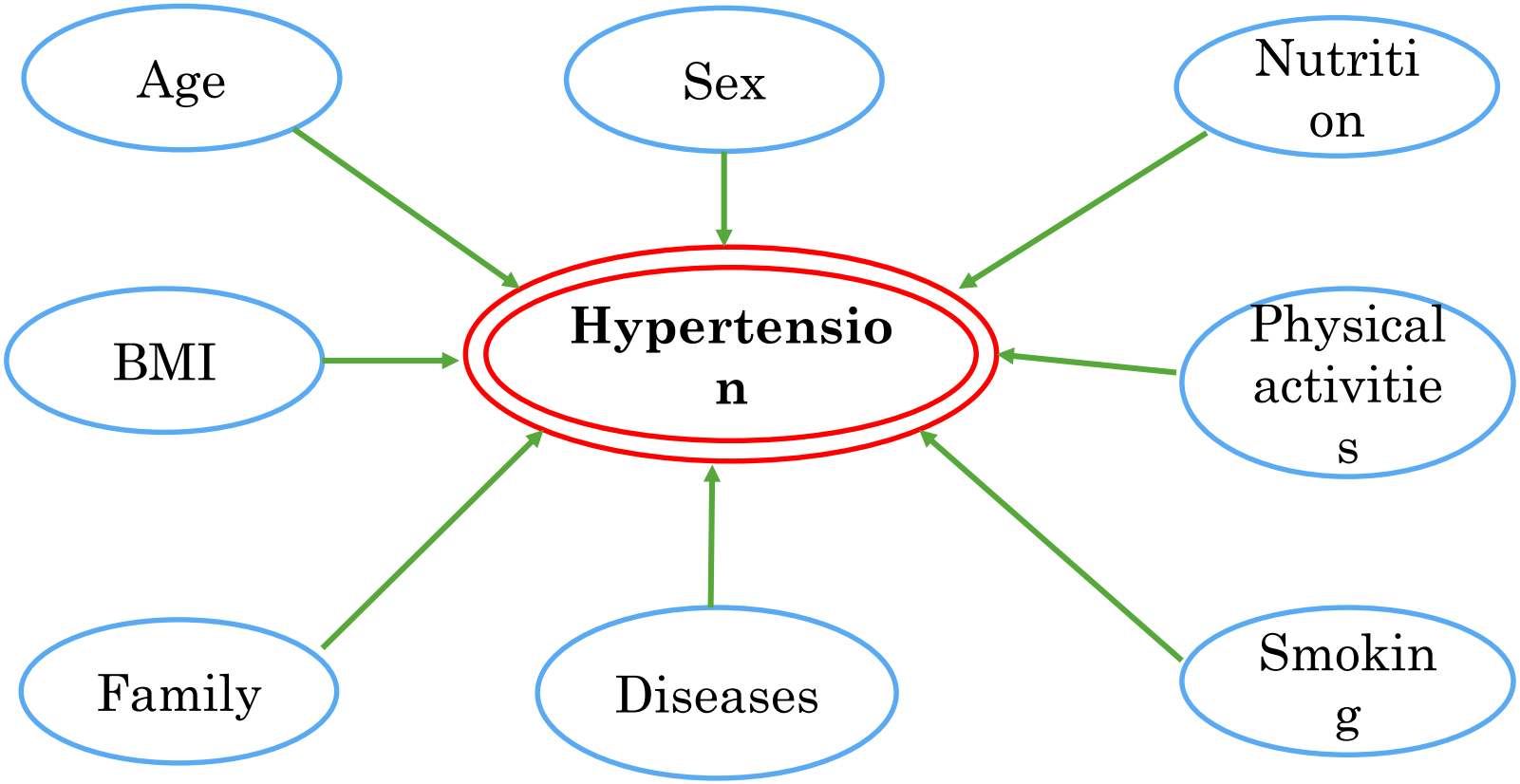
talebi_ghane@yahoo.com

نقش متغیرها

۱. **متغیر وابسته:** متغیری که متغیر مستقل بر روی آن اثر می‌گذارد. به عبارت دیگر مساله همان متغیر وابسته است زیرا بر اثر یک یا چند عامل بوجود آمده است.

۲. **متغیر مستقل**: متغیری است که محقق تاثیر آن را بر سایر متغیرها مورد سنجش قرار می‌دهد. به بیانی دیگرعوامل احتمالی ایجاد آن را متغیر مستقل گویند.

talebi_ghane@yahoo.com

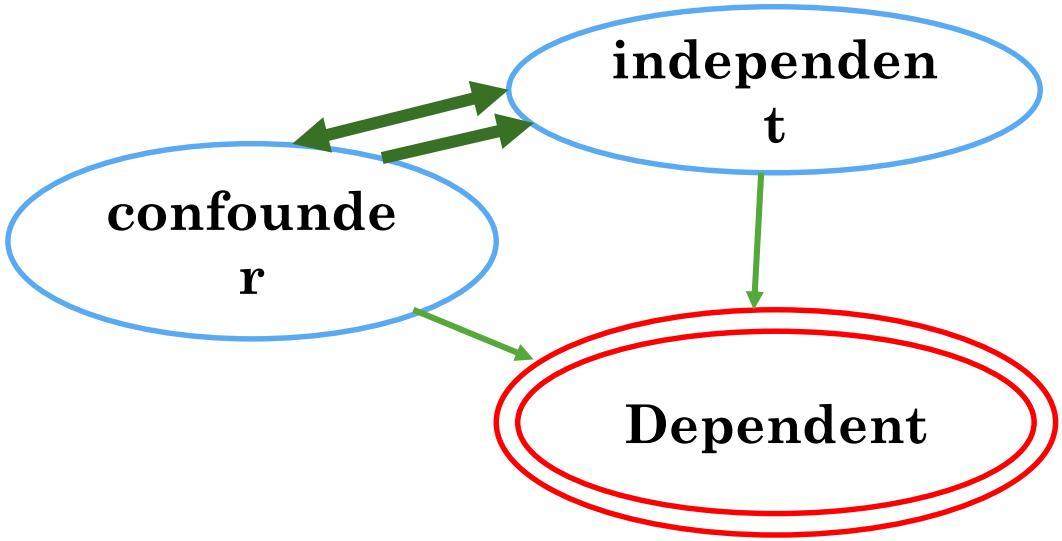


talebi_ghane@yahoo.com

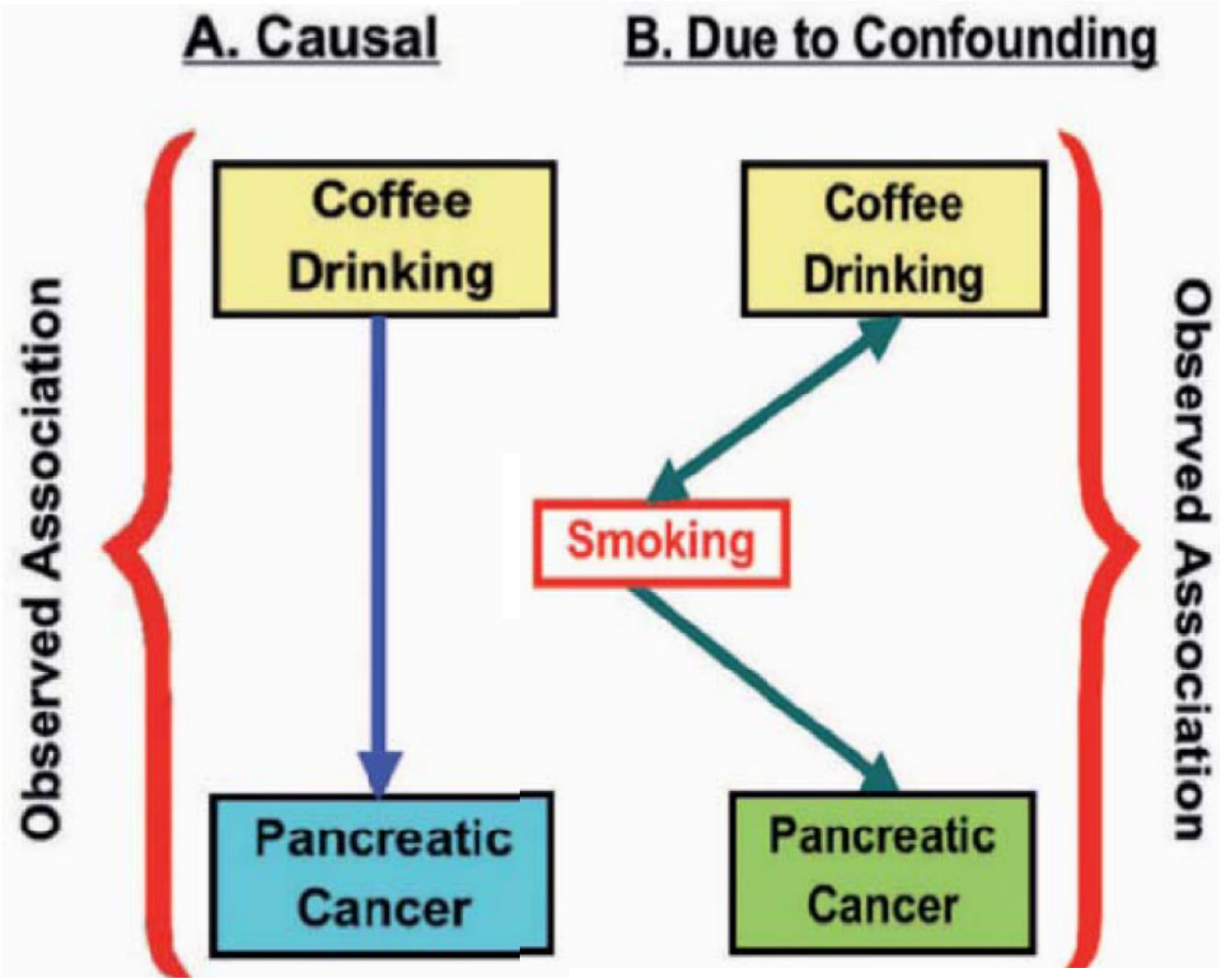
نقش متغیرها

- **متغیر زمینه ای :** حالت عمومی دارند و در همه تحقیقات وجود دارند (سن وجنس)

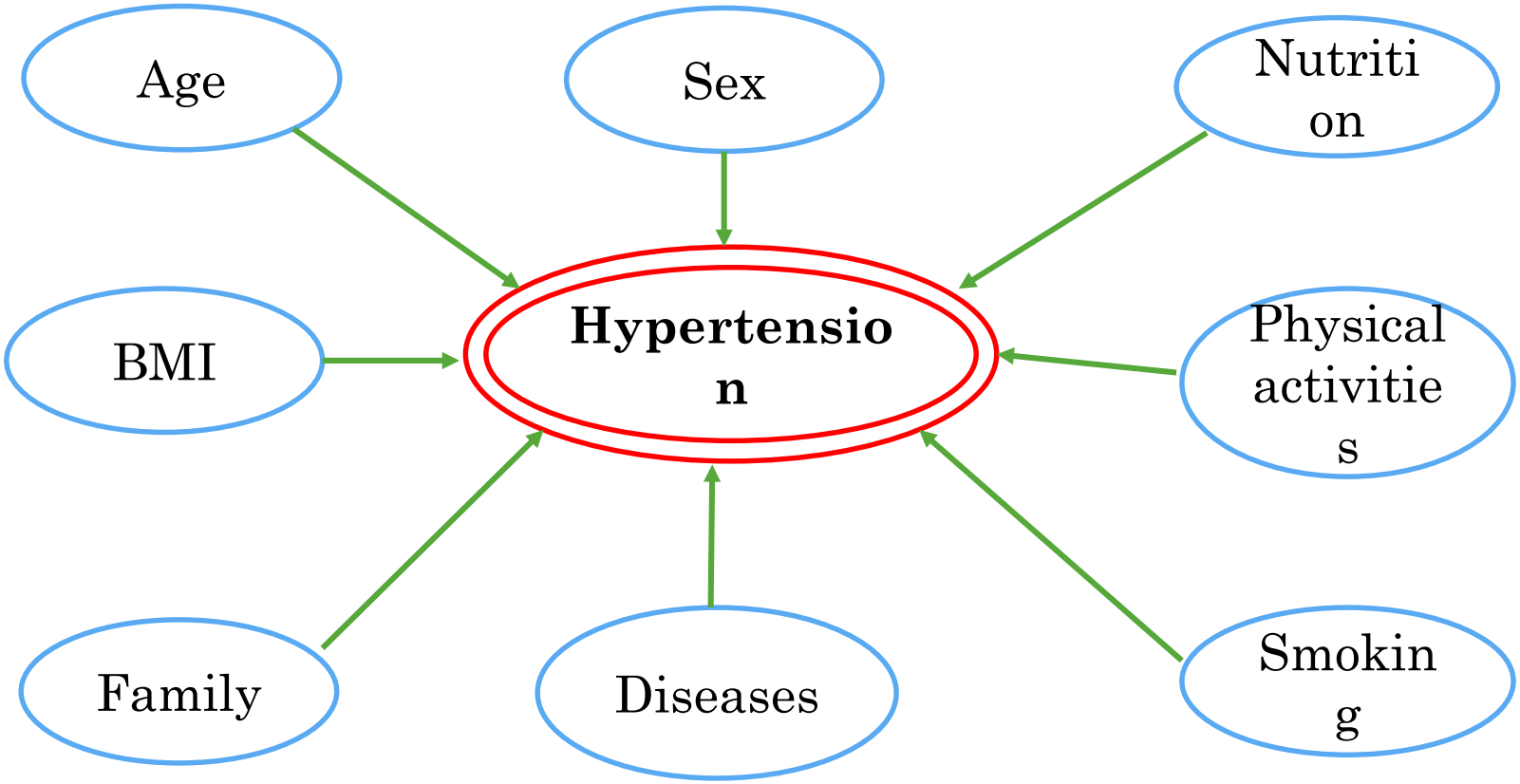
- **متغیر مداخله گر(مخدوش کننده) :** متغیری است که بر روی رابطه علت معلولی بین متغیر مستقل و متغیر وابسته تاثیر می‌گذارد و باعث قوی یا ضعیف شدن کاذب رابطه بین آن‌ها می‌شود.



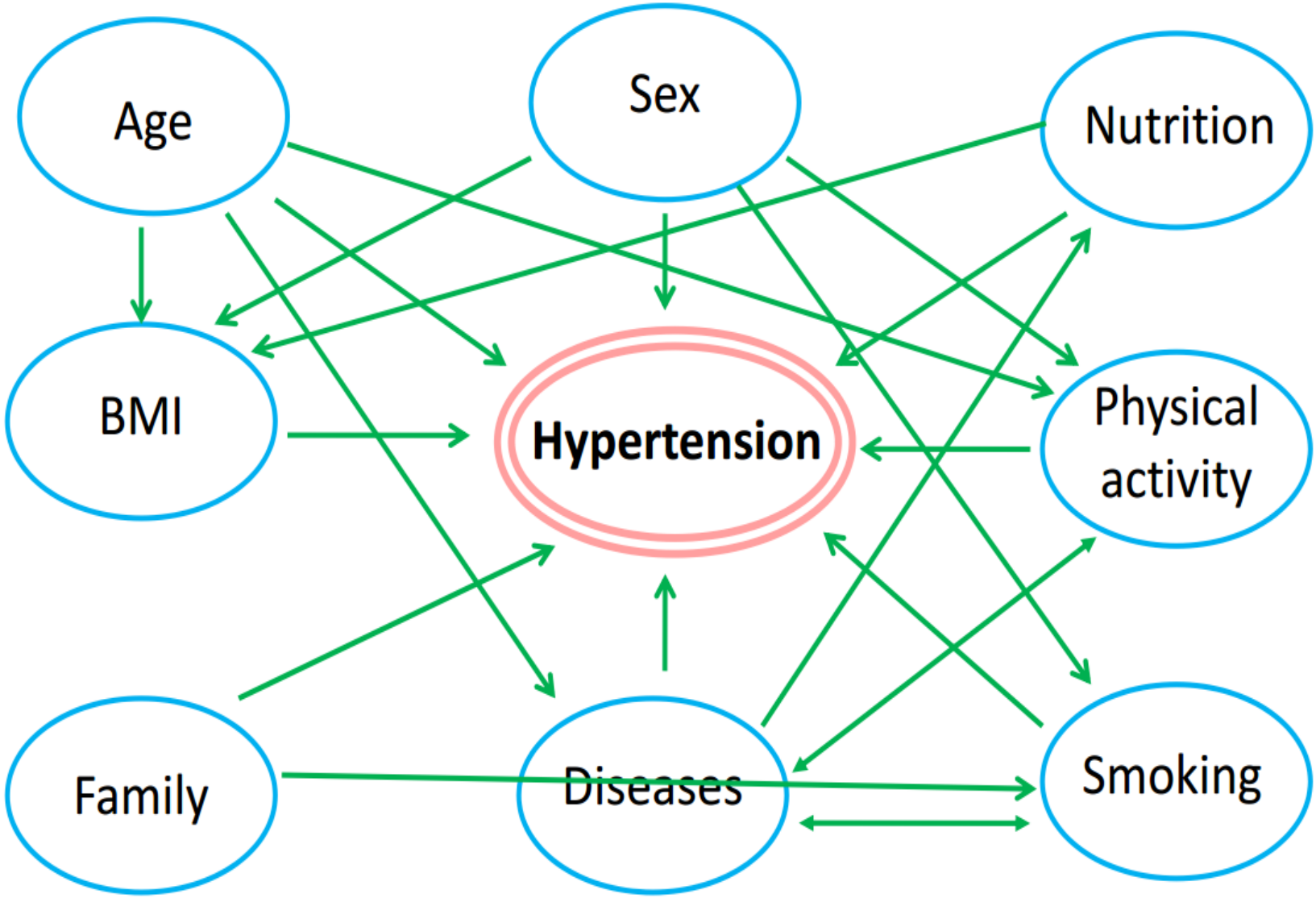
talebi_ghane@yahoo.com



talebi_ghane@yahoo.com



talebi_ghane@yahoo.com


Age
Sex
Nutrition
BMI
Hypertension
Physical activity
Family
Diseases
Smoking

روشهای آماری برای تجزیه و تحلیل نتایج

talebi_ghane@yahoo.com

علم آمار

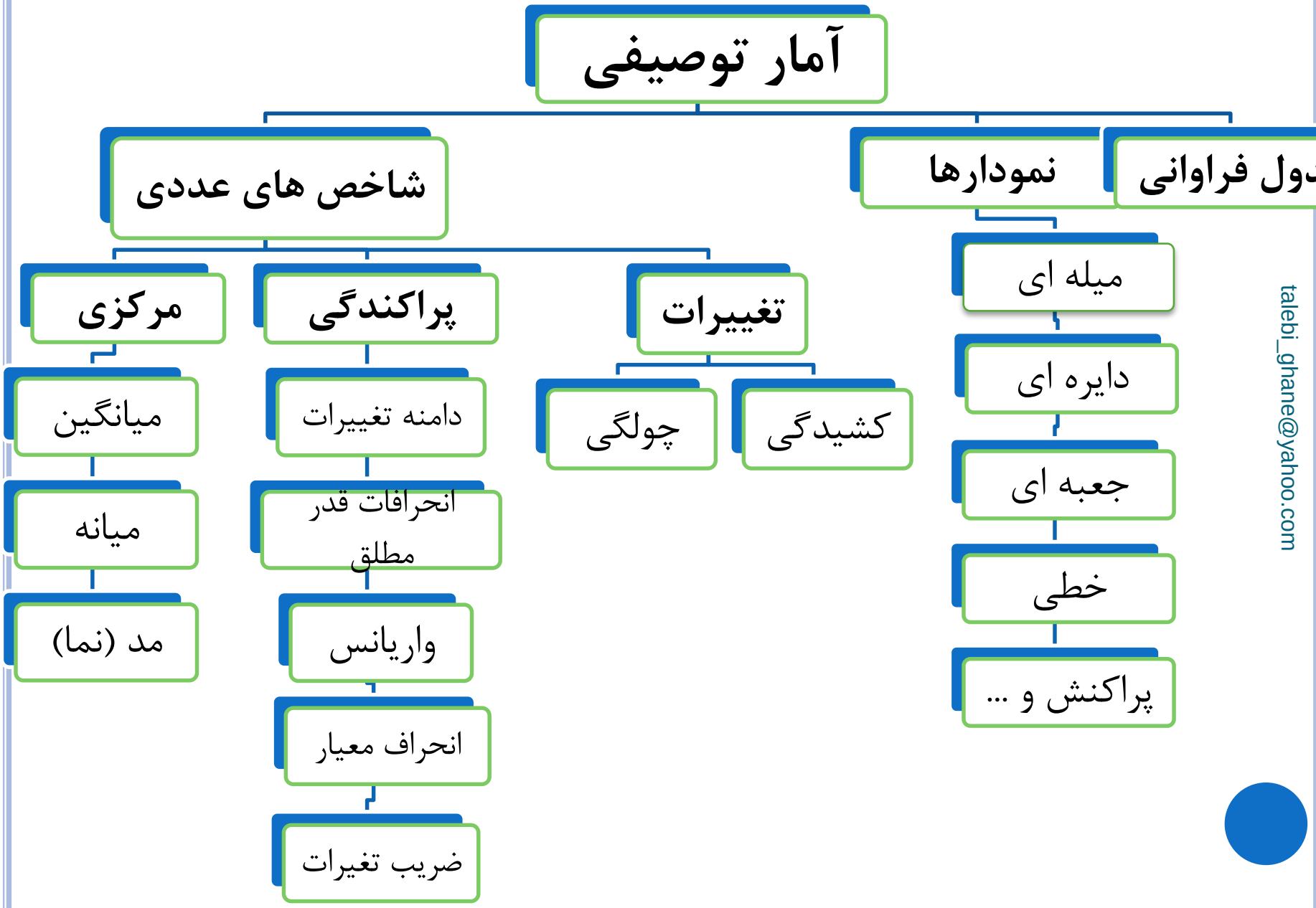
روشهای آماری به دو قسمت اصلی تقسیم می شود:

- **آمار توصیفی**:
  - جمع آوری داده ها
  - توصیف داده ها
    - جداول فراوانی
    - نمودارها
    - گزارش شاخصهای عددی (شاخصهای مرکزی و پراکندگی)

- **آمار استنباطی**: استنباط از داده های نمونه با استفاده از برآورد و آزمون های آماری و تعمیم نتایج حاصل از نمونه به کل جامعه.

آمار توصیفی

talebi_ghane@yahoo.com



talebi_ghane@yahoo.com

آمار استنباطی (تحلیلی)

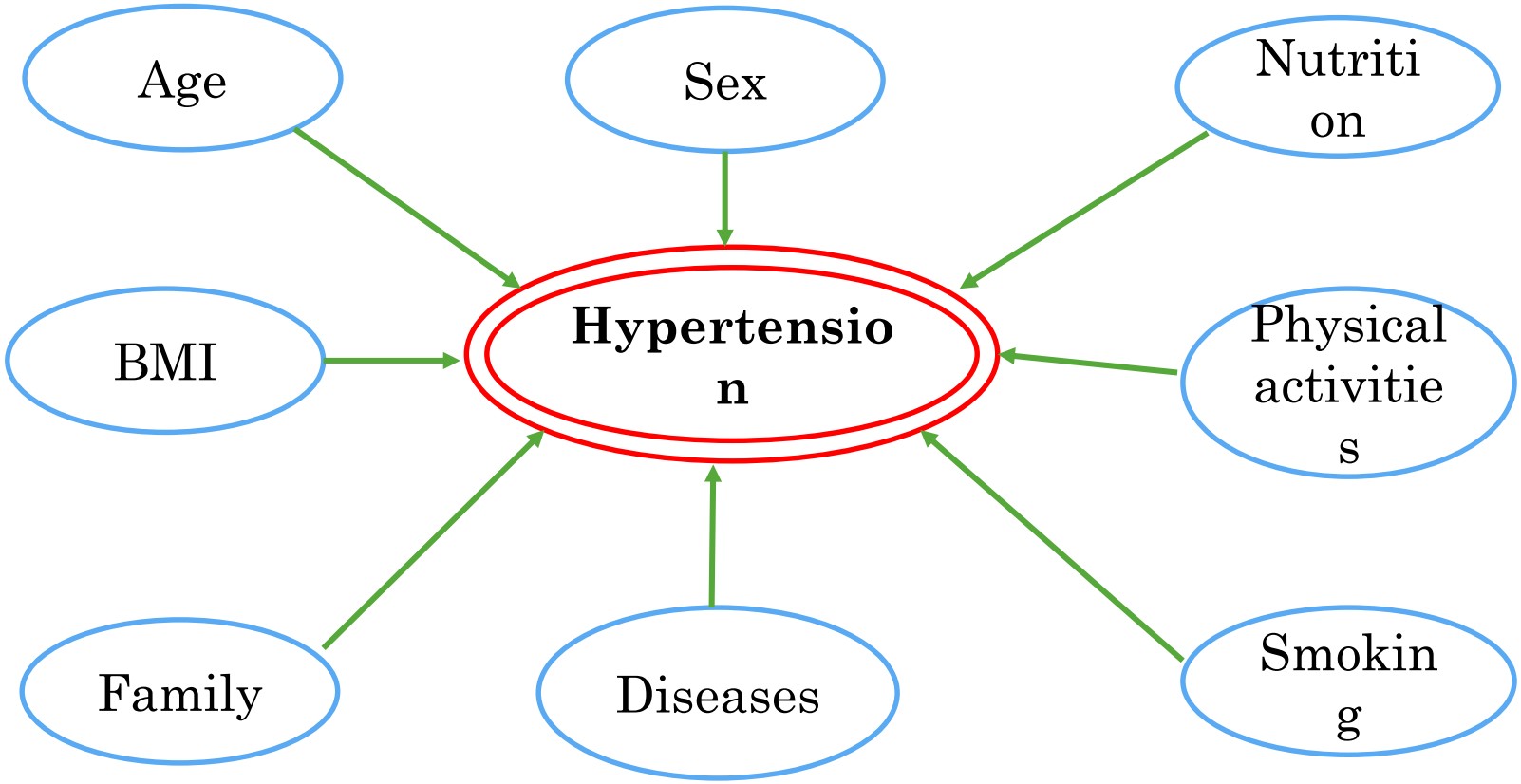
talebi_ghane@yahoo.com

ملاحظات انتخاب آزمون آماری

- متغیر وابسته (پیامد) کمی می‌باشد یا کیفی (رتبه ای یا اسمی)؟
- تعداد گروه‌های مورد مقایسه چه تعداد است؟
  - دو گروه
  - بیش از دو گروه
- آیا گروه‌های مورد مقایسه مستقل هستند یا وابسته؟
- آیا توزیع داده‌ها نرمال است یا خیر؟
  - آزمونهای پارامتری (با فرض نرمال بودن داده ها)
  - آزمونهای ناپارامتری (با فرض نرمال نبودن داده ها یا رتبه ای بودن متغیر)

- استفاده از آزمون **کلموگروف اسمیرنوف یا شپیروویلک** برای بررسی نرمال بودن متغیرها
- چک کردن همگنی واریانسها (آزمون لون و ...)

آزمون فرضیه برای یک جامعه



talebi_ghane@yahoo.com

آزمون فرضیه برای یک جامعه

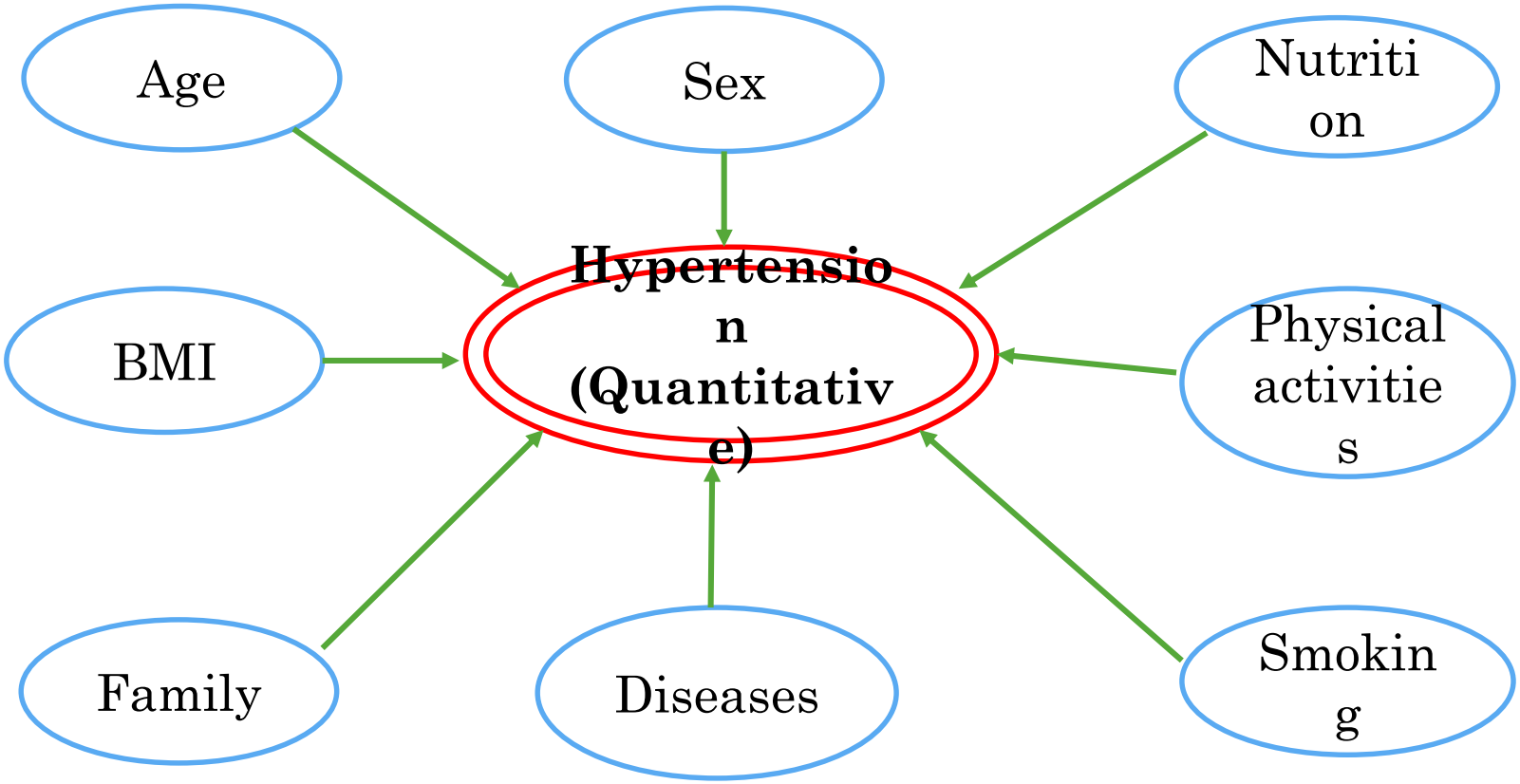
- میانگین فشار خون سیستولیک افراد در جامعه برابر ۱۴ است.
- شیوع سیگاریها در بیماران مبتلا به پرفشاری خون در جامعه کمتر از ۱۰٪ است.
- میانگین نمره تغذیه بیماران مبتلا به پرفشاری خون بیش از ۷۰ است.
- One Sample T-test

talebi_ghane@yahoo.com

بررسی ارتباط دو متغیر

talebi_ghane@yahoo.com

آزمون فرضیه برای بررسی ارتباط بین دو متغیر



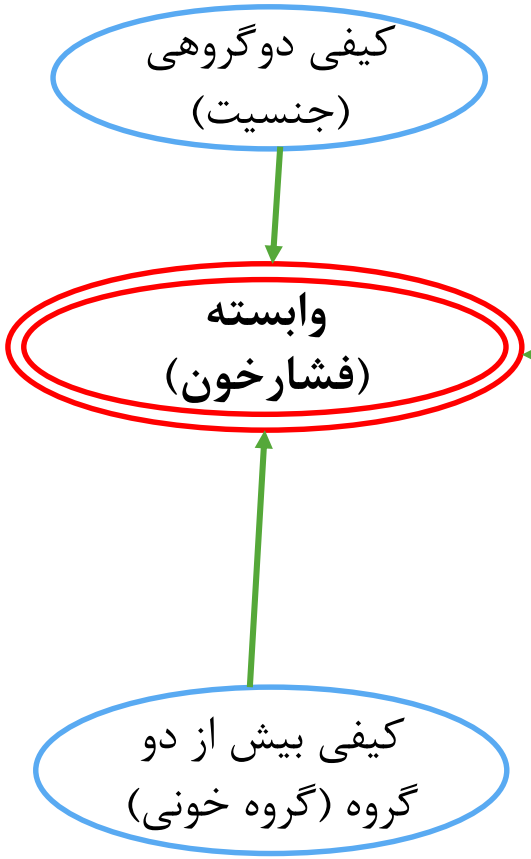
talebi_ghane@yahoo.com

بررسی ارتباط یک متغیر کمی یا رتبه ای با یک متغیر کیفی مستقل

- بررسی نرمال بودن متغیر وابسته در هر زیرگروه
- مقایسه میانگین متغیر کمی در دو گروه مستقل

۱- متغیر کیفی دوگروهی:

- آزمون پارامتری: t-test مستقل
- آزمون ناپارامتری (عدم نرمال بودن متغیر یا رتبه ای بودن متغیر وابسته): من کیفی (جنسیت، گروه خونی)



- مقایسه میانگین متغیر کمی در بیش از دو گروه مستقل

۲- متغیر کیفی بیش از دوگروه:

- آزمون پارامتری: آنالیزواریانس یکطرفه (ANOVA)
- آزمون ناپارامتری (عدم نرمال بودن متغیر یا رتبه ای بودن متغیر وابسته): کراسکال والیس

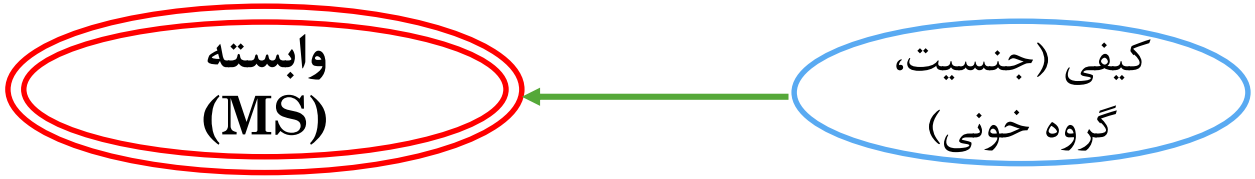
talebi_ghane@yahoo.com

بررسی ارتباط متغیر کمی یا رتبه ای با یک متغیر کیفی مستقل

- جهت مقایسه میانگین در بیش از دو گروه آنالیز واریانس پیشنهاد می‌شود. البته باید توجه داشت که آزمون تحلیل واریانس فقط به این فرضیه پاسخ می‌دهد که آیا میانگین در سه جامعه برابر هست یا خیر؟ حال اگر فرض برابری میانگین‌ها رد شود می‌بایست از آزمون‌های تعقیبی (Post Hoc) استفاده شود تا به‌طور دقیق مشخص شود اختلاف بین کدام گروه‌ها می‌باشد.
- توکی
- بن فرونی
- LSD
- دانت

بررسی ارتباط بین دو متغیر کیفی

- مقایسه نسبت‌ها و یا شیوع در دو یا بیش از دو گروه مستقل
- متغیر کیفی دارای دو یا بیش از دو گروه باشد: کای اسکور



- در استفاده از آزمون کای‌دو هیچ یک از خانه‌های جدول متقاطع نباید خالی باشند و نبایستی بیش از ۲۰٪ خانه های جدول فراوانی مورد انتظار کمتر از ۵ داشته باشند. پیشنهاد می‌شود زمانی که حجم نمونه مورد مطالعه کافی نیست و شرایط ذکر شده محقق نمی‌گردد از آزمون دقیق فیشر استفاده شود.

talebi_ghane@yahoo.com

مثال

قسمت اول – خلاصه مشخصات طرح

عنوان طرح (فارسي): **بررسی ارتباط بین سلول‌های T CD161+ با بروز GvHD حاد در دریافت کنندگان پیوند آلوژن سلول‌های بنیادی خونی.**

عنوان طرح (انگليسي):
The Evaluation of Association between CD161+ T cells and Incidence of aGvHD in Allogenic Hematopoietic Stem Cell Transplant Recipients.

مدير اجرايي طرح[۲] :

همکاران:

دانشکده / مرکز تحقیقاتي : مرکز تحقیقات سلولهای بنیادی و سلول درمانی

گروه :

محیط پژوهش : مرکز تحقیقات سلولهای بنیادی و سلول درمانی

مدت اجرا :

مثال

تعیین ارتباط بین میزان سطح پلاسمایی IL-17A در نمونه خون دریافت کنندگان پیوند آلوژن سلول‌های بنیادی خونی با بروز aGvHDو شدت آن در بیماران.

؟

talebi_ghane@yahoo.com

مثال

سطح پلاسمایی IL-17A: پیوسته کمی
بروز aGvHD: کیفی دو حالته (بروز دارد/ ندارد)
شدت aGVHD: (0/I/II/III/IV)

برای مقایسه میانگین سطح پلاسمایی IL-17A در افرادی که aGvHD رخ داده یا نداده در صورت نرمال بودن t-test مستقل

برای مقایسه میانگین سطح پلاسمایی IL-17A در شدت رخداد aGvHD در صورت نرمال بودن ANOVA

talebi_ghane@yahoo.com

مثال

تعیین ارتباط Donor type (Sibling/unrelated/FMT) با شدت aGvHD

؟

talebi_ghane@yahoo.com

مثال

Donor type: کیفی چند حالته
شدت aGVHD: (0/I/II/III/IV)

در صورت وجود تعداد مناسب در هر خانه آزمون کای اسکور و در صورت مناسب نبودن آزمون دقیق فیشر

talebi_ghane@yahoo.com

بررسی ارتباط یک متغیر کمی یا رتبه ای با یک متغیر کیفی به صورت وابسته

- بررسی نرمال بودن اختلاف دو متغیر وابسته

۱

اندازه قبل
(فشارخون قبل
از مصرف دارو)

۱- دو اندازه تکراری یا جور شده

- آزمون پارامتری: t زوجی (paired t-test)

۲

اندازه بعد
(فشارخون بعد
از مصرف دارو)

- آزمون ناپارامتری (عدم نرمال بودن متغیر یا رتبه ای بودن متغیر): ویلکاکسون

talebi_ghane@yahoo.com

بررسی ارتباط متغیر کمی یا رتبه ای با یک متغیر کیفی به صورت وابسته

۲- چند اندازه تکراری

- آزمون پارامتری: آنالیز واریانس با اندازه های تکراری (Repeated measurement)
- آزمون ناپارامتری (عدم نرمال بودن متغیر یا رتبه ای بودن متغیر وابسته): فریدمن

۱	فشارخون قبل از مصرف دارو
۲	فشارخون بلافاصله بعد از مصرف دارو
۳	فشارخون یک ساعت بعد از مصرف دارو
۴	فشارخون ۲۴ ساعت بعد از مصرف دارو

talebi_ghane@yahoo.com

بررسی ارتباط بین دو متغیر کیفی به صورت وابسته

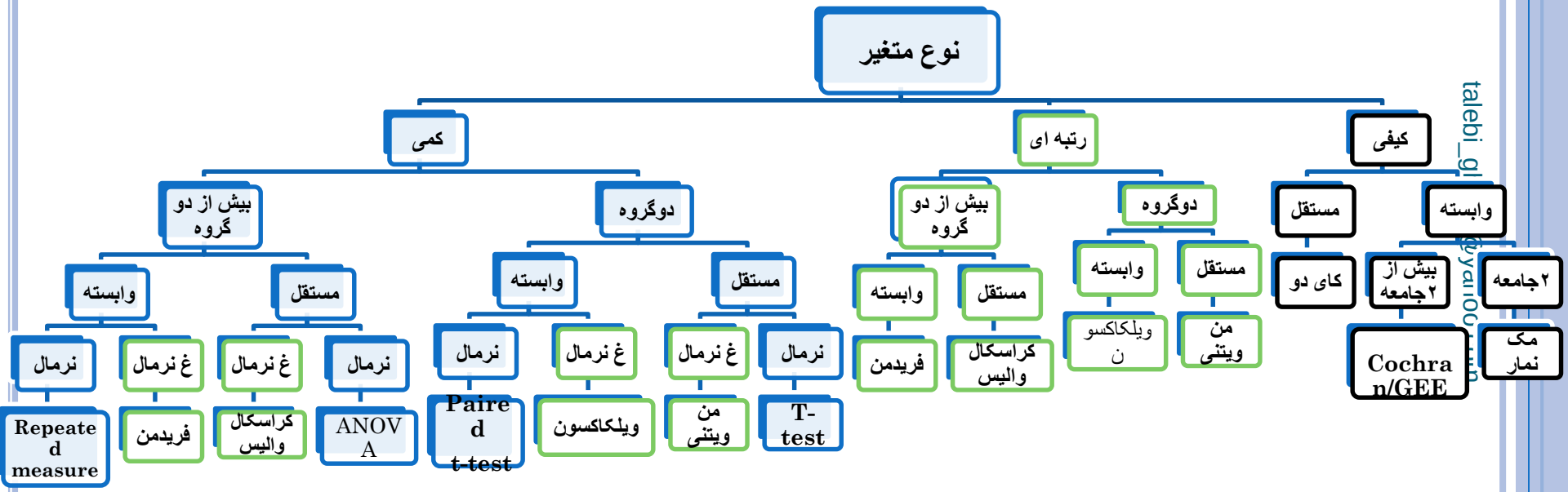
۱- متغیر کیفی دوحالته باشد و دو اندازه تکراری داشته باشد: مک نمار

۲- متغیر کیفی دوحالته باشد و چند اندازه تکراری داشته باشد : کوکران

۳- متغیر کیفی دو یا چند حالته باشد و چند اندازه تکراری داشته باشد : GEE

talebi_ghane@yahoo.com

آزمون های مورد استفاده در بررسی ارتباط بین دو متغیر



بررسی ارتباط بین دومتغیر کمی

- ضریب همبستگی می‌تواند از ۱+ تا ۱- باشد، ۱+ نشان‌دهنده ارتباط کامل و مستقیم و ۱- نشان‌دهنده ارتباط کامل و معکوس می‌باشد.
- معمولاً ضریب همبستگی کمتر از ۰,۲۵ را همبستگی خیلی کم، ۰,۲۵-۰,۵ را همبستگی کم، ۰,۵-۰,۶۹ را متوسط، ۰,۷-۰,۹ را بالا و ۰,۹-۱ را خیلی بالا قلمداد می‌کنند.
- توجه شود که صفر بودن ضریب همبستگی تنها عدم وجود رابطه‌ی خطی بین دو متغیر را نشان می‌دهد ولی نمی‌توان مستقل بودن دو متغیر را نیز نتیجه گرفت.
- همبستگی فقط نشان دهنده ارتباط بین دو متغیر است و نمی‌تواند هیچ ادعایی از نظر ارتباط علی بین دو متغیر داشته باشد.

talebi_ghane@yahoo.com

بررسی ارتباط بین دو متغیر کمی

- ضریب همبستگی پیرسون (محاسبه میزان همبستگی دو گروه متغیر کمی نرمال مستقل)
- ضریب همبستگی اسپیرمن (محاسبه میزان همبستگی دو گروه متغیر کمی غیر نرمال مستقل یا رتبه ای)
- ضریب تطابق کندال (محاسبه میزان تطابق دو گروه متغیر رتبه ای وابسته)

talebi_ghane@yahoo.com

بررسی ارتباط بیش از دو متغیر

talebi_ghane@yahoo.com

آنالیز واریانس دو یا سه طرفه

- گاهاً پیش می‌آید که تاثیر دو یا سه متغیر کیفی بر یک متغیر کمی سنجیده می شود.در این شرایط از آنالیز واریانس دوطرفه یا سه طرفه استفاده می شود.
- مقایسه دو داروی جدید با داروی قبلی با درنظر گرفتن دوزهای مختلف و تاثیر آن بر برخی بیومارکرها.

تحلیل کوواریانس

- گاهاً پیش می‌آید که متغیری به نام متغیر همراه وجود دارد که اثر آن بر روی متغیر وابسته شناخته شده است و عدم برخورد با آن باعث خطا در اندازه گیری خواهد شد.
- تصور کنید شما از دو رویکرد آموزشی متفاوت در دو کلاس درس استفاده می‌کنید و در پایان نمره آنها را با هم مقایسه می‌کنید. حال اگر سطح اولیه دانش دو کلاس درس در خصوص موضوع مطرح شده در ابتدای مداخله با هم متفاوت باشد، اختلاف سطح اولیه دانش بین دانش آموزان دو کلاس، مقایسه ما را با خطا مواجه می‌کند.
- برای کنترل کردن متغیر همراه که در این مثال سطح اولیه دانش بود از روش آنالیز کوواریانس استفاده می‌شود.
- متغیر همراه الزاماً می‌بایست کمی باشد.

مثال

- تاثیر امواج مغناطیسی بر IL6, IL21 بعد از تحریک سیستم ایمنی
- امواج مغناطیسی با ۴ طول موج مختلف

؟

مثال

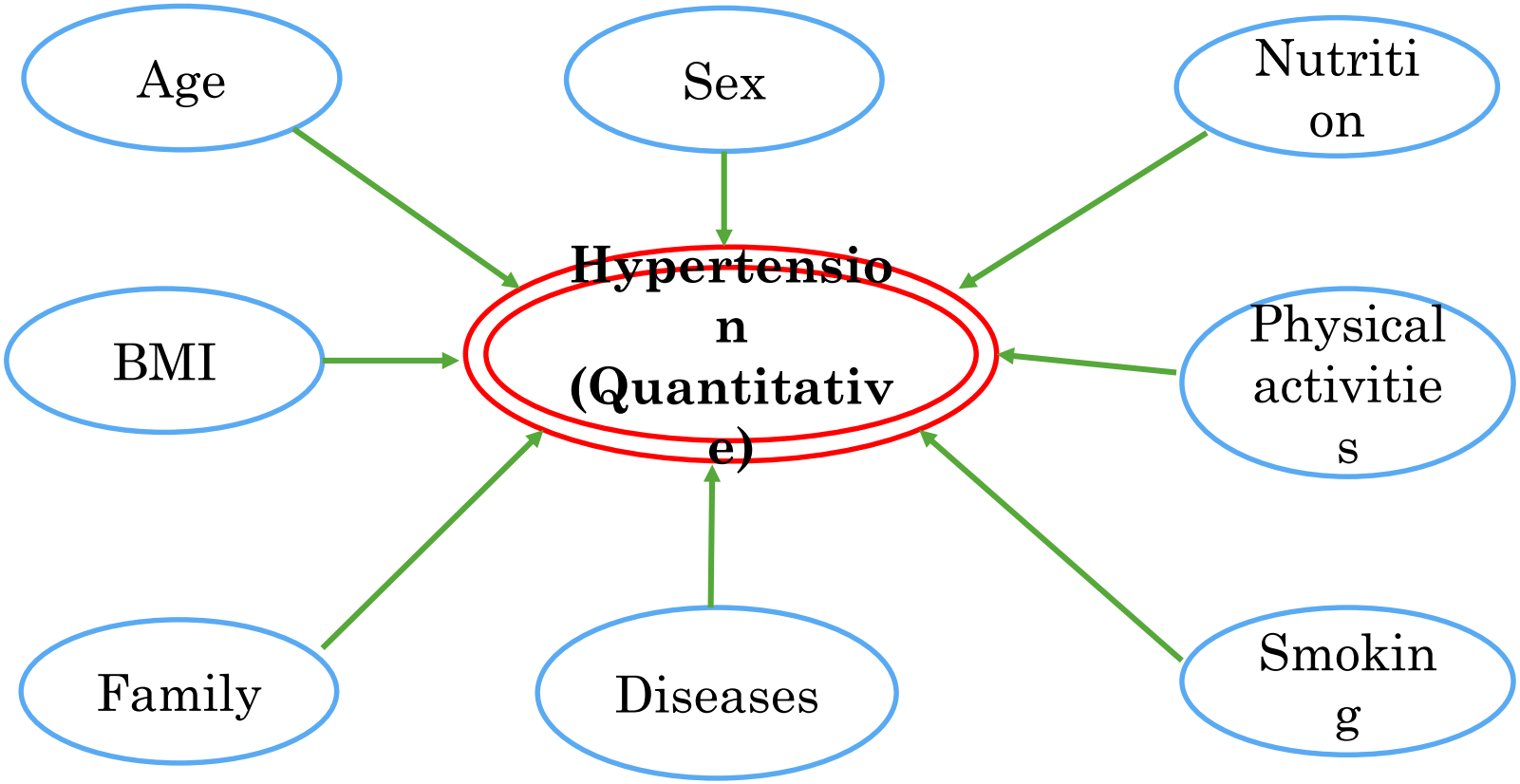
- ANOVA
- نمره IL6, IL21 قبل از تحریک سیستم ایمنی هم ثبت شده.
- آیا نمره IL6, IL21 قبل از تحریک سیستم ایمنی در ۴ گروه امواج مغناطیسی یکسان بوده و این نمره بر نمره بعد از تحریک تاثیر نمی گذارد؟

- استفاده از ANCOVA معقولتر است.

بررسی تاثیر همزمان تمامی متغیرها بر متغیر وابسته

talebi_ghane@yahoo.com

بررسی تاثیرهمزمان متغیرهای مختلف بر متغیر وابسته



talebi_ghane@yahoo.com

رگرسیون

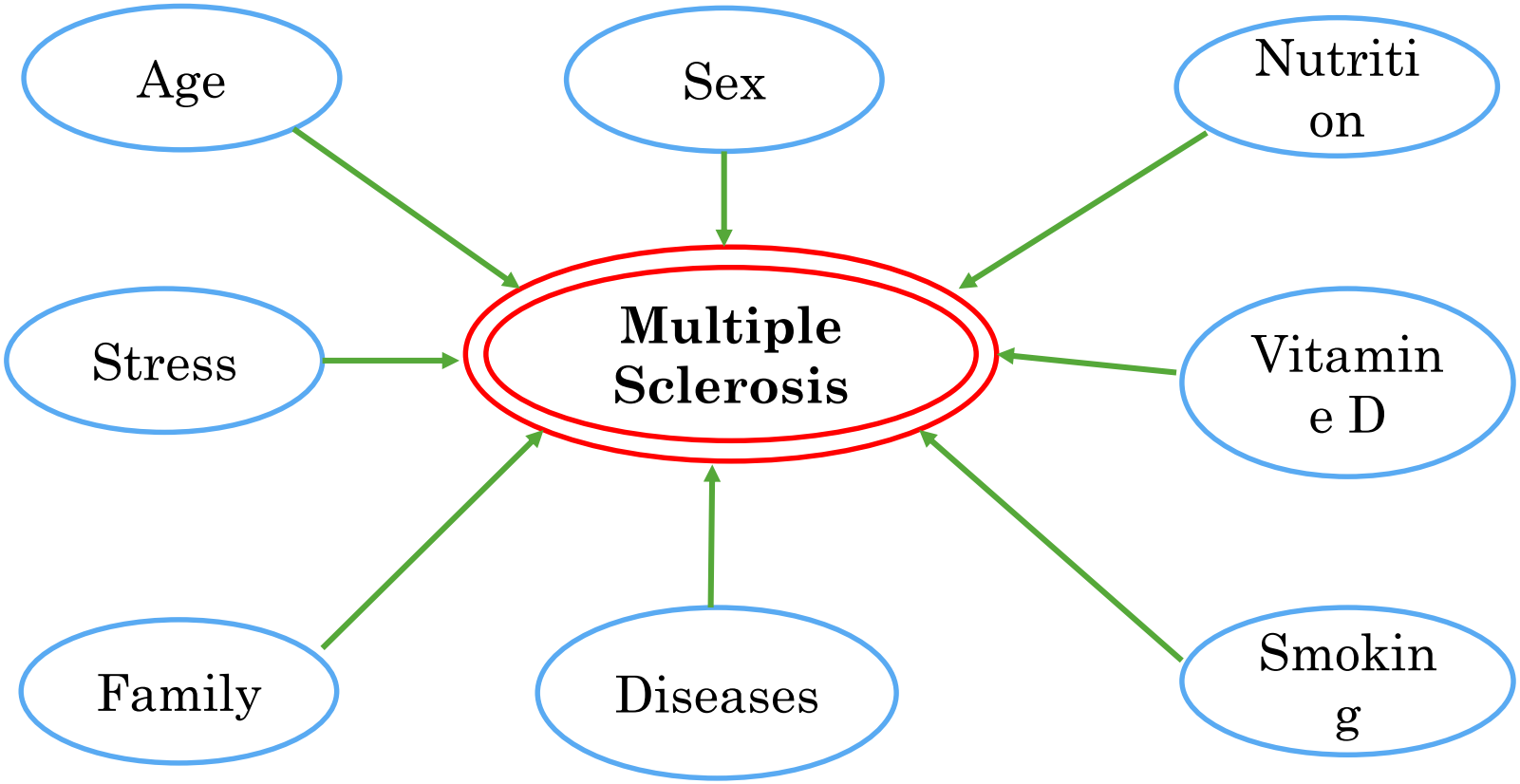
- در تحلیل‌های پزشکی معمولاً به منظور پیش‌بینی و یا بیان تغییرات یک متغیر بر اساس اطلاعات متغیر دیگر، از روش‌های رگرسیونی استفاده می‌شود.
- مدل رگرسیون خطی: اگر برای شناسایی و پیش‌بینی متغیر وابسته کمی فقط از یک متغیر مستقل استفاده شود، مدل را رگرسیون خطی ساده می‌گویند.

$$Y=\beta 0+\beta 1X+\epsilon$$

اگر تعداد متغیرهای مستقل دو متغیر باشند معادله رگرسیونی به شکل زیر نوشته می‌شود و به آن رگرسیون چندگانه گویند:

$$Y=\beta 0+\beta 1X1+\ \beta 2X2+\ \epsilon$$

بررسی تاثیرهمزمان متغیرهای مختلف بر متغیر وابسته



talebi_ghane@yahoo.com

رگرسیون

- مدل رگرسیون لجستیک: اگر بخواهیم رابطه بین یک متغیر مستقل را با یک متغیر وابسته با مقدارهای کیفی دو حالته بسنجیم، در این حالت روش عادی رگرسیون خطی جوابگو نخواهد بود و باید از رگرسیون لجستیک استفاده کرد. مثلاً در خصوص رابطه بین ساعات ورزش روزانه و ابتلا به دیابت، متغیر وابسته ما یعنی ابتلا به دیابت دو حالتی است و فرد یا به دیابت مبتلا هست و یا خیر لذا در این مثال روش رگرسیون لجستیک کاربرد دارد.
- مدل رگرسیون لجستیک چندسطحی: اگر بخواهیم رابطه بین یک متغیر مستقل را با یک متغیر وابسته با مقدارهای کیفی چندحالته بسنجیم از این مدل استفاده می شود.
- برای تحلیل داده‌های حاصل از شمارش مدل رگرسیونی پواسون استفاده می-شود.
- مدل رگرسیونی کاکس کاربردی‌ترین مدل در تحلیل بقا می‌باشد.

با تشکر از حسن توجه شما