IGF GENE-NETWORK IDENTIFIES SPECIFIC INTERACTIONS UNDERLYING EXTREME SBP IN DEVELOPING CHILDREN

NZSA Conference
University of Canterbury
November 2015

P.G. Parmar,

C.E. Pennell, L.J. Palmer, L. Briollais

AUT UNIVERSITY

Department of Biostatistics and Epidemiology
School of Public Health and Psychosocial Studies
Faculty of Health and Environmental Sciences

OUTLINE

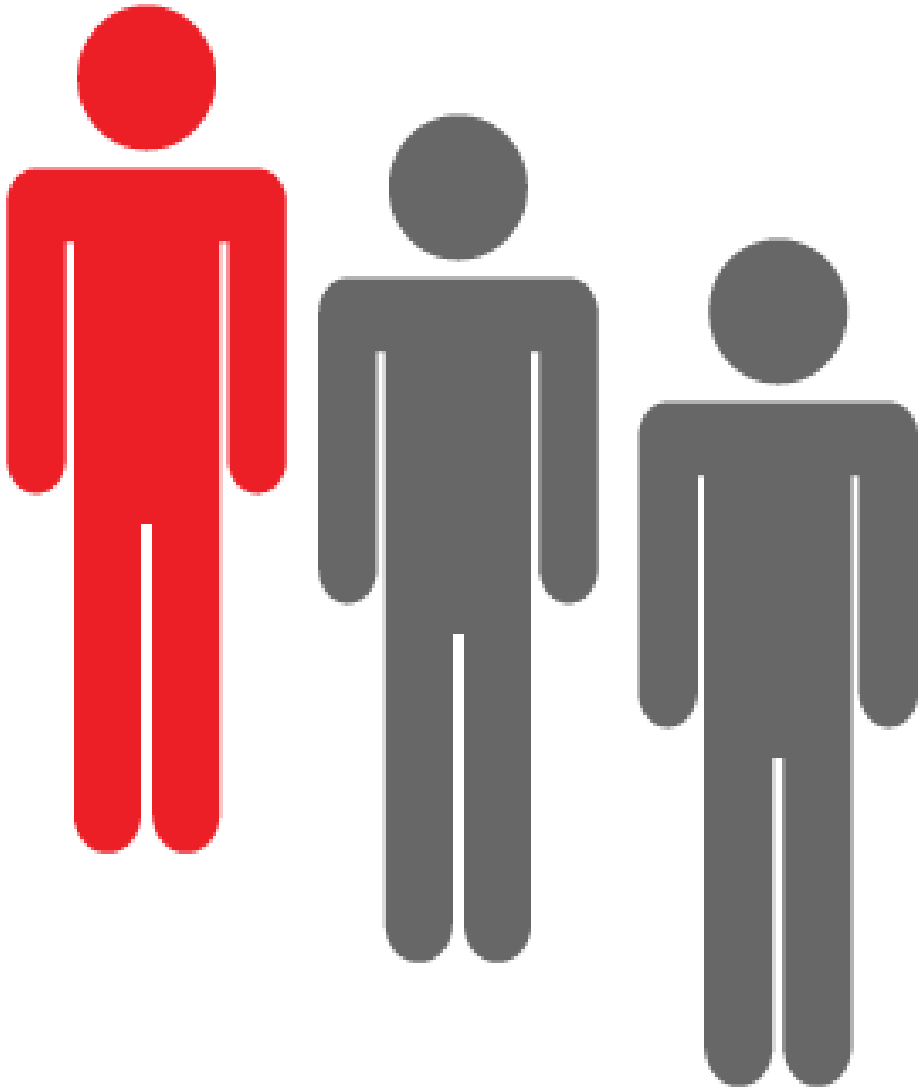
- Aim
- Background
- Study design
- Analyses
- Results
- Conclusions
- Future research

AIM

Investigate if modelling the insulin-like growth factor (IGF)-Axis as a gene-network can identify gene-interactions predictive of the early onset of elevated systolic blood pressure (SBP) in developing children.

RATIONALE

- Elevated BP is a symptom and a disease
- Consistent, persistent elevated BP (140mmHg SBP/90mmHg DBP) = Clinical Hypertension
  - Established cardiovascular (CVD) risk-factor
  - Is the leading cause of adult death and disability in the world through its predisposition to CVD events e.g. MI, stroke, obesity, type-II diabetes, metabolic syndrome

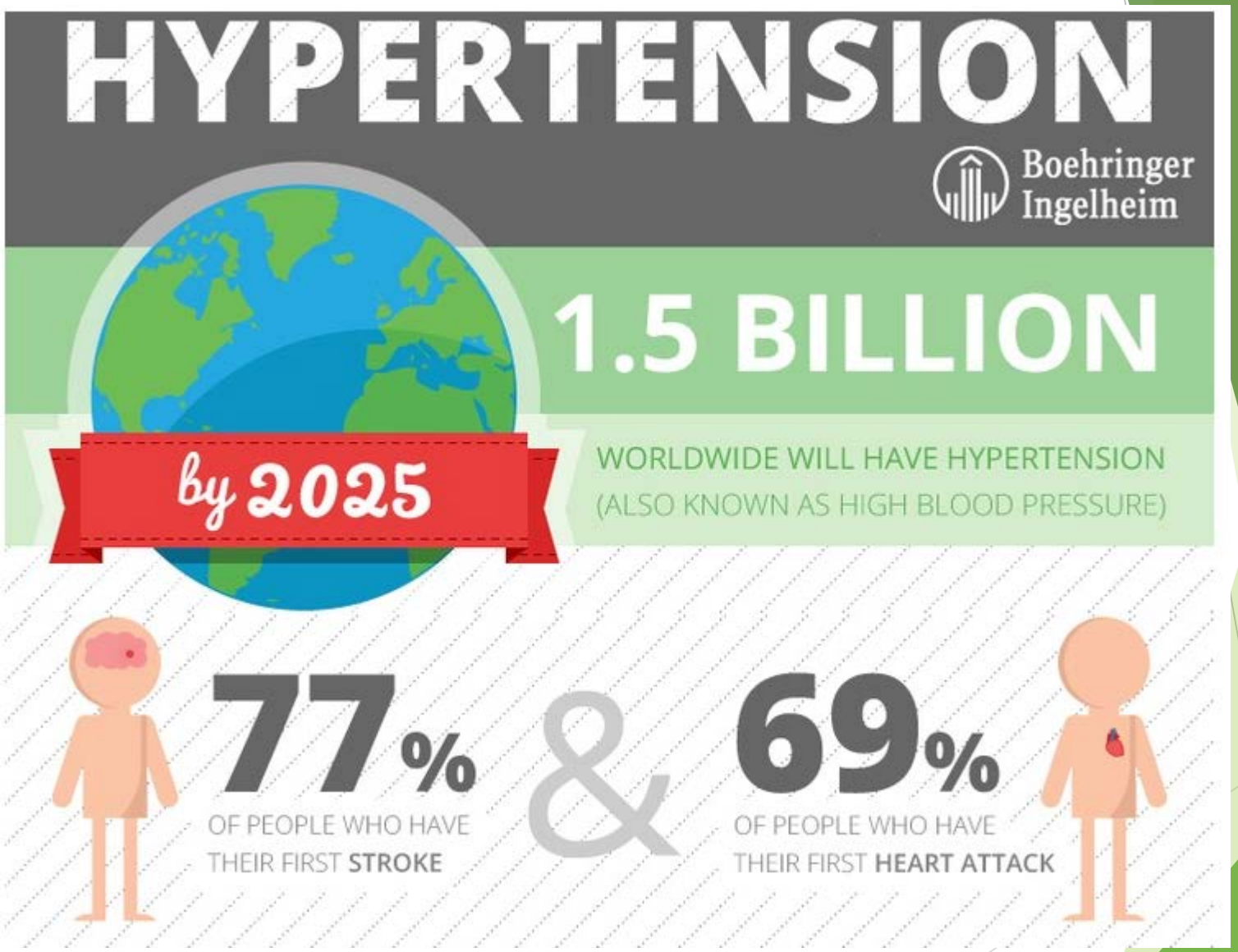


NEARLY 1 IN 3

PEOPLE HAS HIGH BLOOD PRESSURE.

Source: National Heart, Lung, and Blood Institute

ConsumerReports®

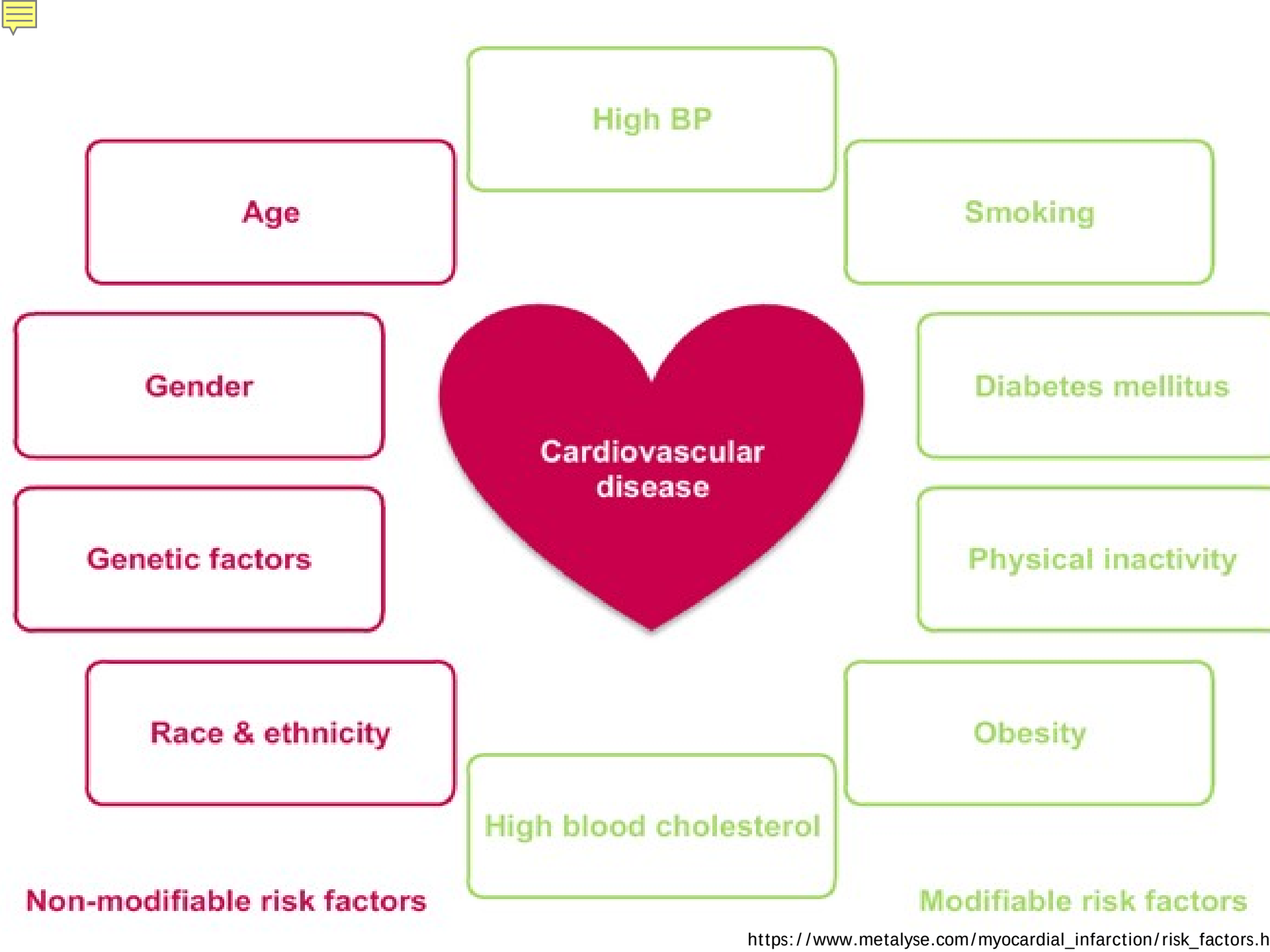

HYPERTENSION
Boehringer Ingelheim
1.5 BILLION
by 2025
WORLDWIDE WILL HAVE HYPERTENSION
(ALSO KNOWN AS HIGH BLOOD PRESSURE)
77%
OF PEOPLE WHO HAVE THEIR FIRST STROKE
&
69%
OF PEOPLE WHO HAVE THEIR FIRST HEART ATTACK

Currently

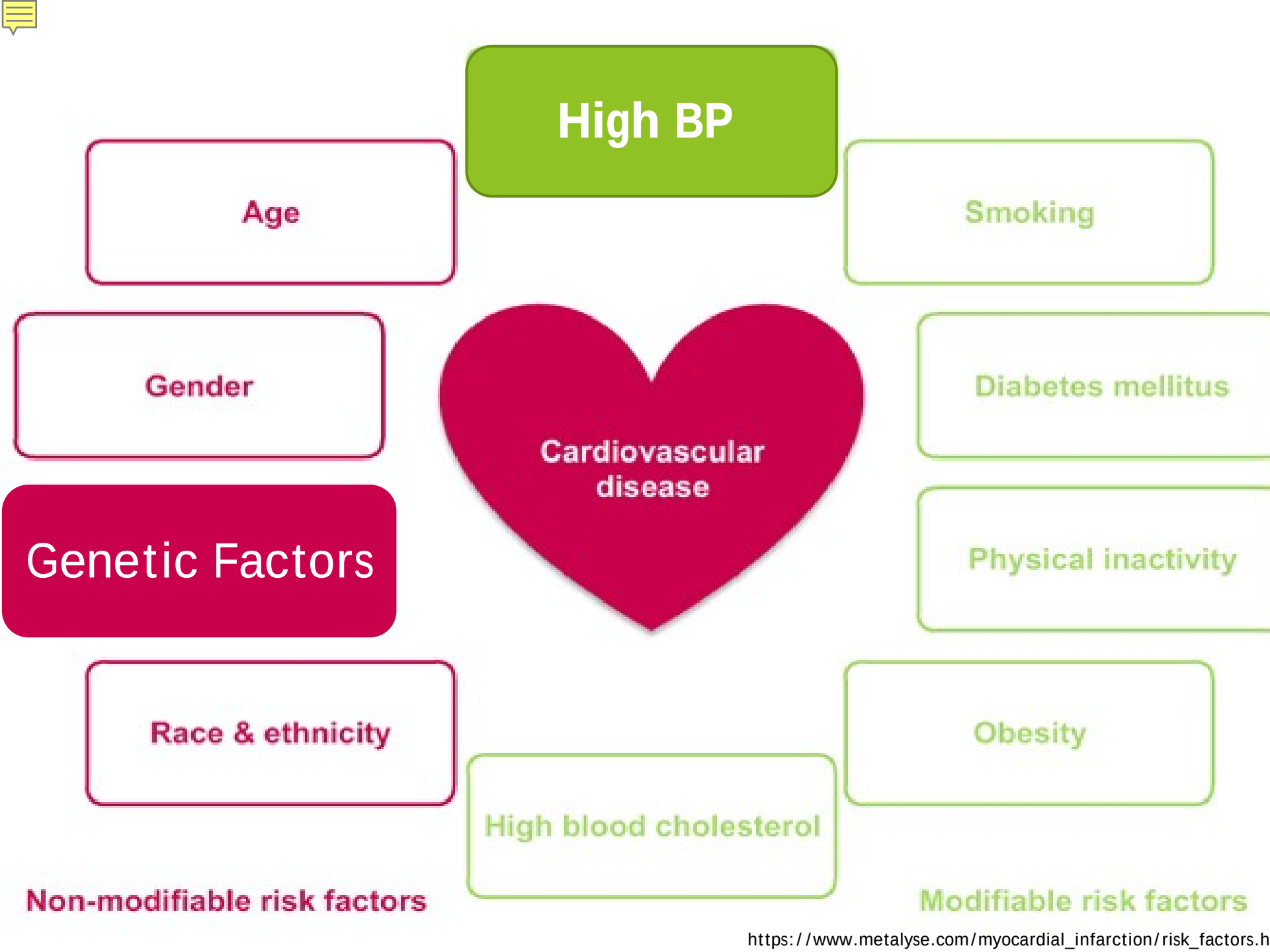
IT ACCOUNTS FOR

7.5 MILLION

DEATHS PER YEAR

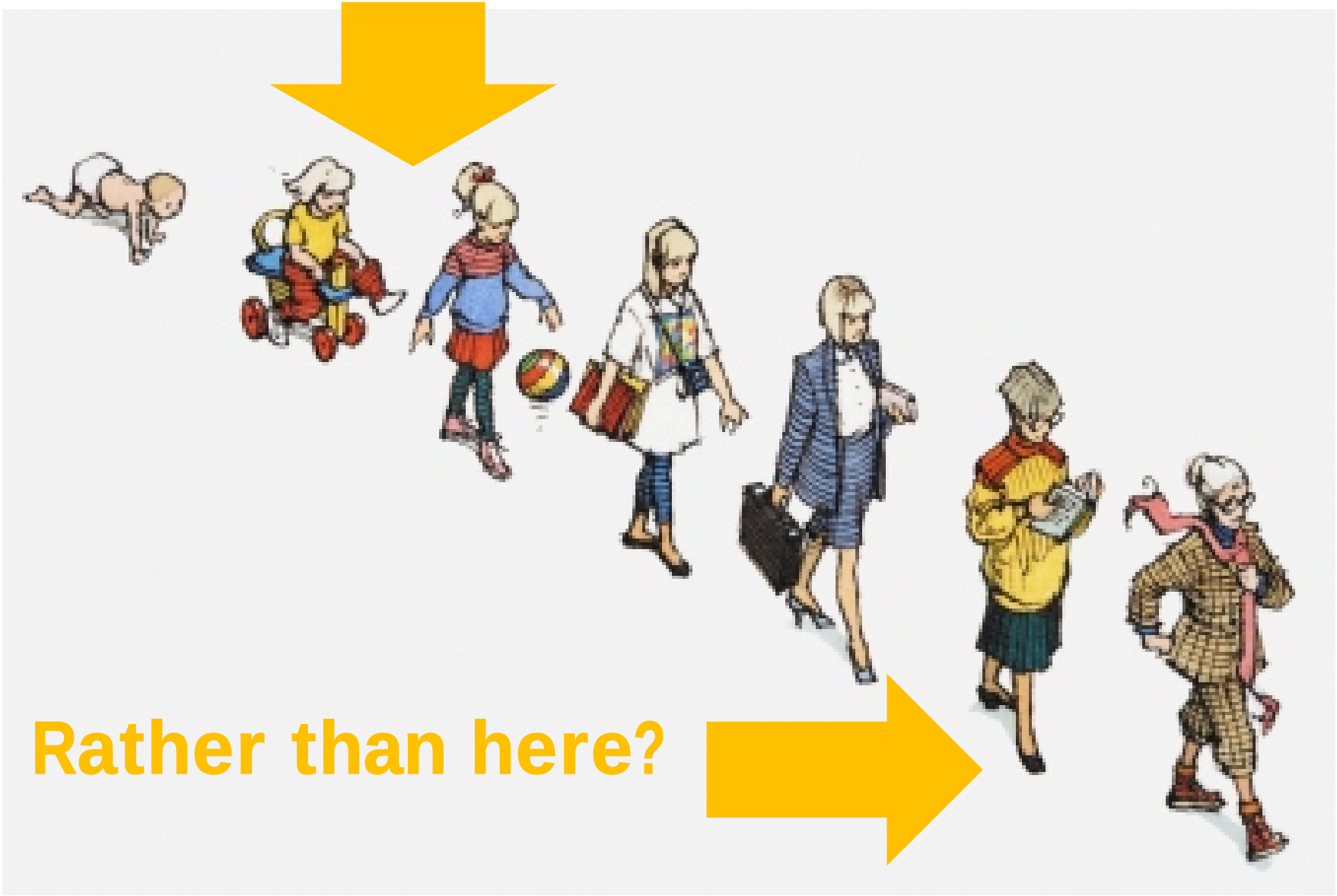


https://www.metalyse.com/myocardial_infarction/risk_factors.h

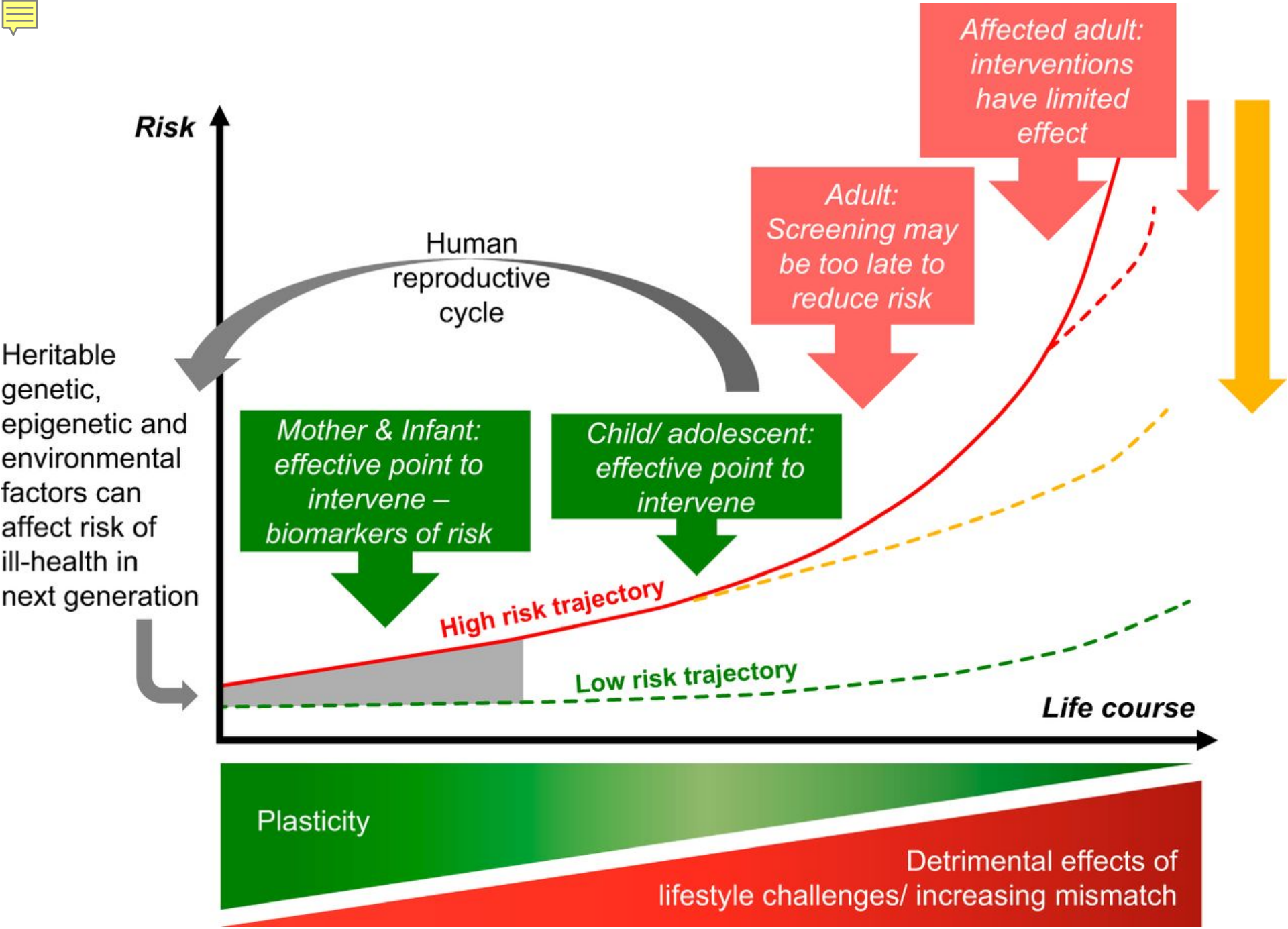


https://www.metalyse.com/myocardial_infarction/risk_factors.h

Why are we looking here?



Rather than here?



Hanson MA, Gluckman PD. Early developmental conditioning of later health and disease: physiology or pathophysiology? Physiol Rev. 2014 Oct;94(4):1027-76.

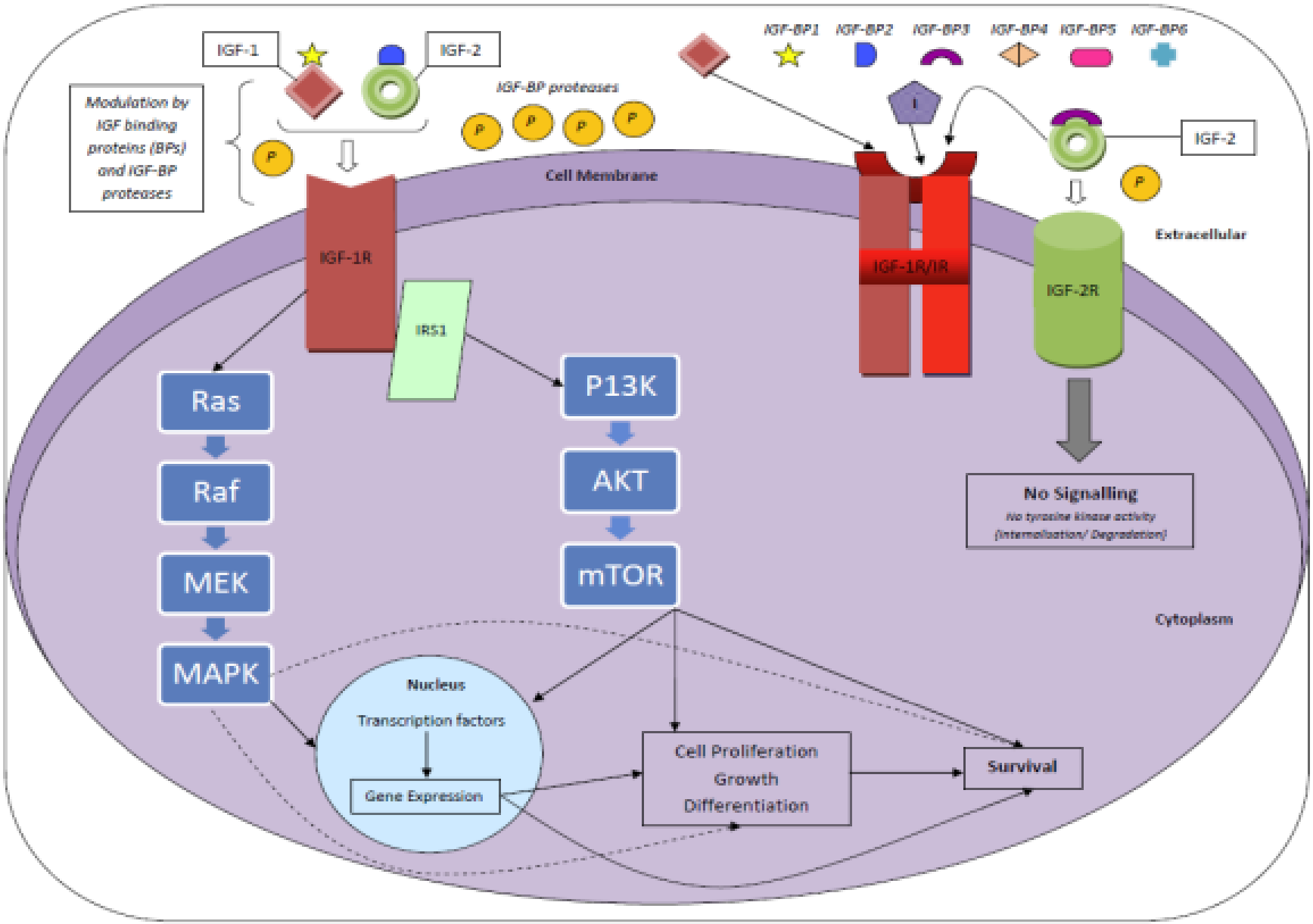
BP DATA

- Raine (Western Australian Longitudinal Pregnancy and Birth Cohort)

- Data from 2,279 individuals who had at least one SBP measure recorded at 5, 8, 10, 12, 14 and 17 years of age

- The most being recorded at age five (1,947 individuals) and the least at the 17 year follow-up (1,249 individuals)

GENETIC DATA: IGF-Axis



IGF-Axis

Gene	Number of SNPs	Percentage of SNPs
IGF-1	15	9
IGF-1R	82	50
IGF-2	8	5
IGF-2R	38	23
IGFBP1	5	3
IGFBP2	4	2
IGFBP3	5	3
IGFBP4	3	2
IGFBP5	5	3
Total	165	100

THE MODEL

- Following Tsonaka *et al*, we used a two-stage mixed effects model to estimate a gene interaction network involving the nine genes of the IGF axis and model their association with the longitudinal SBP outcome

- Stage 1 generates gene-based random effects summarising the SNP effects within each gene

- Stage 2 uses the gene-specific random effects from stage 1 to model the gene-gene interactions against SBP

STAGE 1: Genetic Model

- Here we model the probability of carrying a particular SNP allele (the rare allele) as a function of fixed and random effects that represent genetic variations of the SNP genotypes at *S* loci studied
  - A logistic model is used to specify the probability to carry at least one rare allele (dominant model) or two rare alleles (recessive model) as a Bernouilli trial
- This model will return the empirical Bays (EB) estimates of the random effects

STAGE 1: Genetic Model

Let π_{ij} be the probability for an individual i $(i = 1, \dots, n)$ to carry a variant in the SNP $j = (1, \dots, 165)$ across the particular gene it belongs to, given we have 9 genes $g = (1, \dots, 9)$. The gene-based lme can be written as:

$$\log \frac{\pi_{ij}}{1-\pi_{ij}} = X'_{ij}\beta + b_i + b_{ij} + b_{ijg}$$

- β' is a vector of regression coefficients on specific covariates $X_{ij}'s$ that could account for specific SNP characteristics, if any is included.
- $b_{i0} \sim N(0, \sigma^2)$ is a random intercept effect for each individual i.
- $b_{ij} \sim N(0, \sigma_g^2)$ is a random effect for each SNP j within individual i. It is assumed independent for each individual i and independent of b_{i0}, but the b_{ij}' s are correlated across genes.
- $b_{ijg} \sim N(0, \sigma_g^2)$ is a random effect for each gene g and SNP j within individual i. It is assumed independent for each individual i and independent of $b_i \; and \; b_{ij}$.
- For each individual and each gene, an empirical bayes estimate can be obtained from the gene-based lme by adding the three random effects: $\widehat{eb_{ig}} = \hat{b}_i + \hat{b}_{ij} + \hat{b}_{ijg}$

STAGE 2: Longitudinal Model

- Here gene-specific estimates are tested using the EB estimates $\left(\widehat{eb_{ig}^{*}}\right)$ from the first stage as covariates in the mixed-effects model for longitudinal SBP
- To model the longitudinal outcome, a linear mixed-effects model with random intercepts and random slopes was used
- A model fitting procedure was implemented to identify the optimal set of variables that should be included in this longitudinal gene-network model
  - Age, Sex and BMI

STAGE 2: Longitudinal Model

$$SBP_{it} \sim X'_{it}\beta + \sum_{g=1}^{G}\sum_{g'>g}^{G}\{\widehat{eb^{*}_{ig}}\} * \{\widehat{eb^{*}_{ig'}}\}\gamma$$

$$+ \sum_{g=1}^{G}\sum_{g'>g}^{G}\{\widehat{eb^{*}_{ig}}\} * \{\widehat{eb^{*}_{ig'}}\} * (BMI_{it} + (Age_{it} * Sex_i))\eta + u_i + \varepsilon_{it}$$

- t $(= 1, \ldots, 5)$ represents the number of follow-up visits at approximately (1) 5 years, (2), 8 years, (3) 10 years, (4) 14 years and (5) 17 years
- X_{it} = Design matrix for the fixed effect covariates which include; BMI, sex and age
- Age_{it} = 5-14 years vs. 15+ years. In brief, these timeframes were chosen to reflect dramatic changes in developmental growth. Further justification surrounding these 'epochs' can be found in chapters six and seven.
- Sex_i = Males or Females
- BMI_i = per unit (kg/m^2) increase in year of age
- β= Vector of regression coefficients on the fixed effects variables BMI, sex and age
- $\widehat{eb^{*}_{ig}}$ = gene-specific empirical Bayes random effects estimated from stage 1
- γ = vector of regression coefficients on the gene-gene specific effects. It quantifies the effect of the pathway components on the longitudinal SBP outcome.
- η = regression coefficients that capture the variation of gene by gene interactions with age (as a discrete variable) and sex on SBP. It quantifies the age-specific effect of the pathway on the longitudinal SBP outcome.
- $u_{i \sim N(0, \sigma^2)}$ = random intercept for each individual i
- $\epsilon_{it} \sim N\ (0, \sigma^2_{\epsilon})$ = Within individual error term with variance σ^2_{ϵ}
- The $u'_i s$ are assumed correlated and independent from the $\varepsilon_i{}'s$.

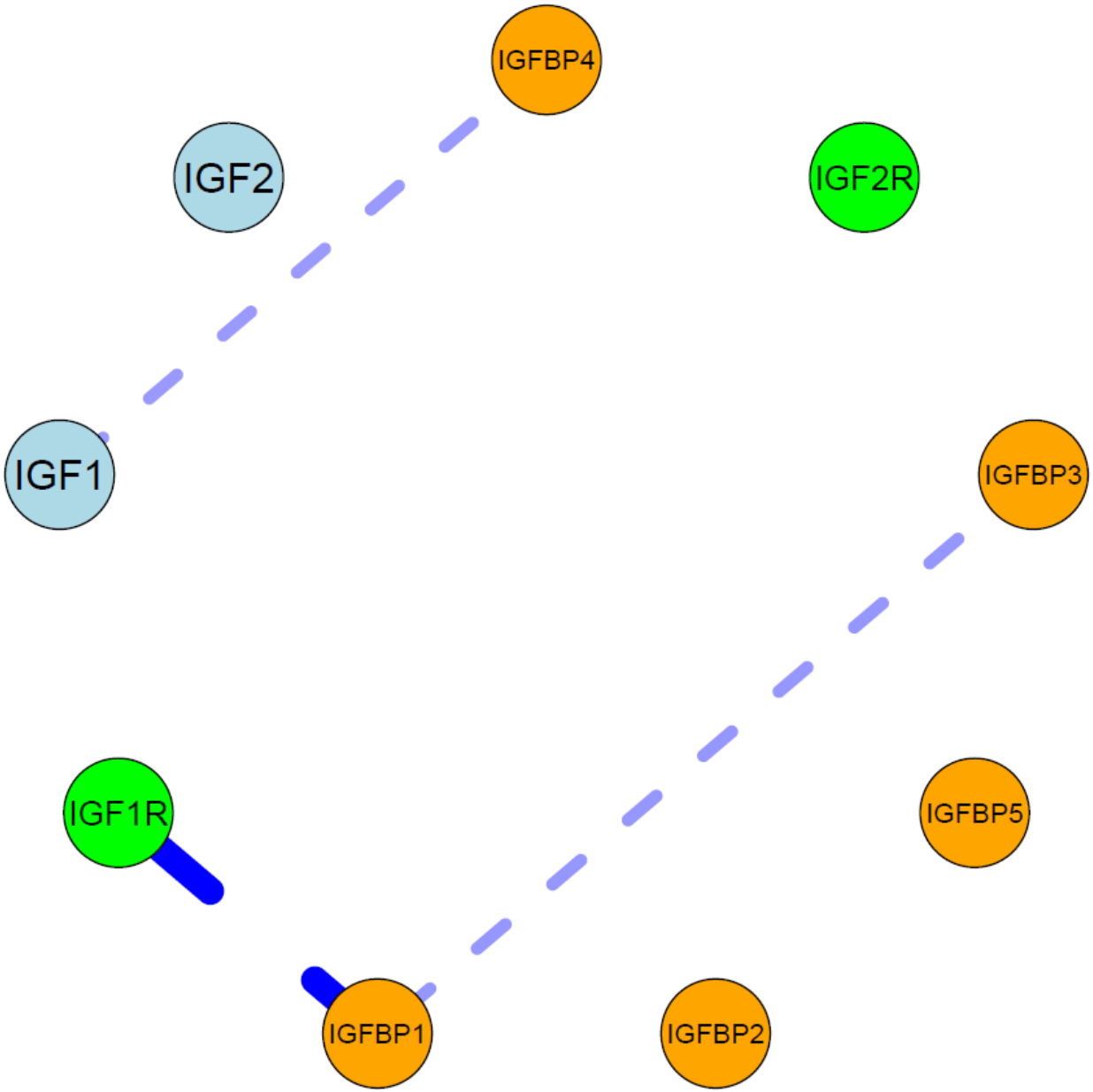
Results

Comparing extreme gene-network profiles

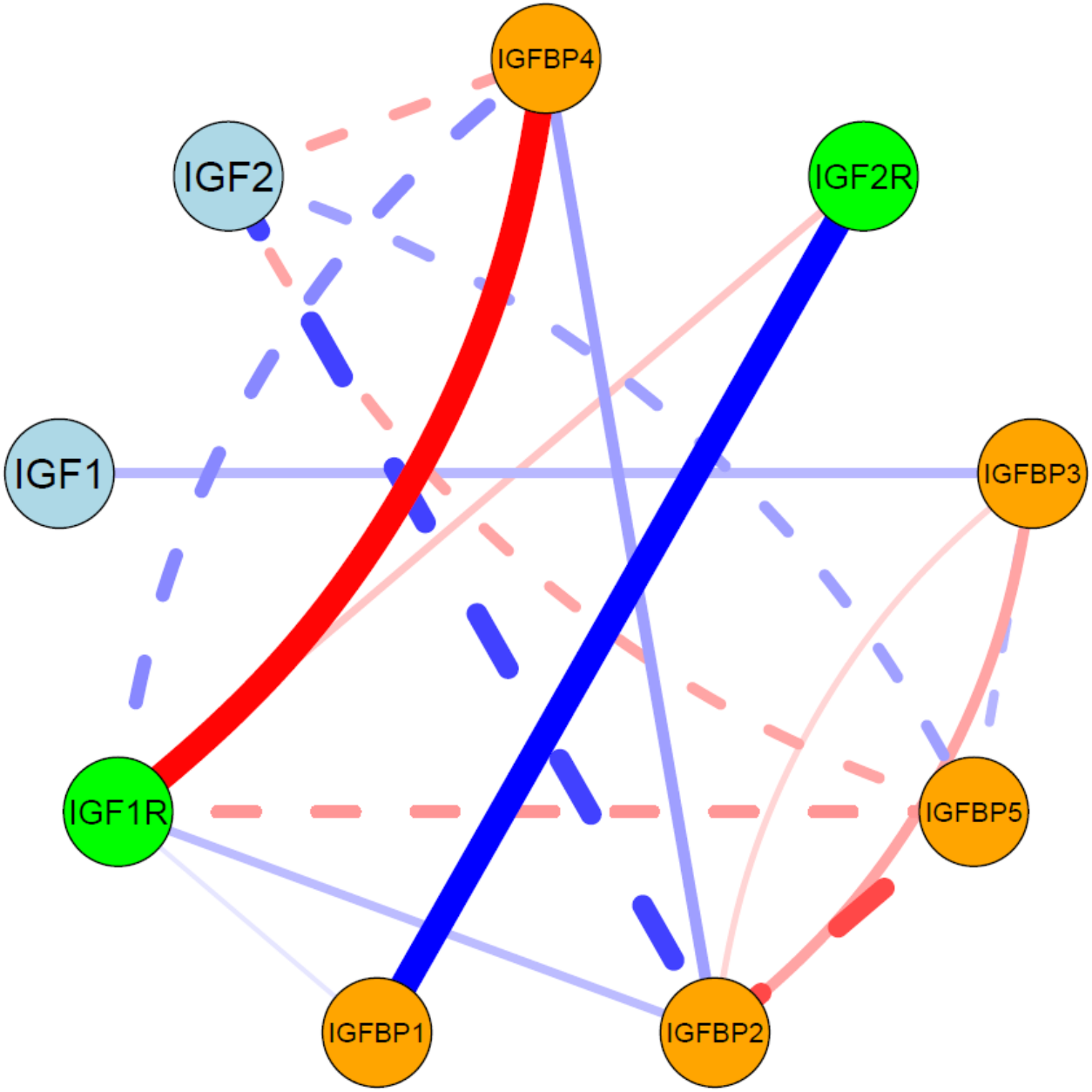
Figure legends

- Only presenting Bonferroni-corrected p-values < 0.0001
- Effects are presented separately by sex
  - Females
  - Males
- Increases to SBP are denoted by dashed lines
- Reductions to SBP are denoted by solid lines
- Thickness of the line indicates the level of significance (i.e. thicker line = more statistically significant than thinner lines)
- Colour of the line indicates the effect size (i.e. the darker the line, the larger the effect compared to lighter coloured lines = smaller the effect size)

Extreme top 5% (< 15 years)



Extreme top 5% (15+ years)

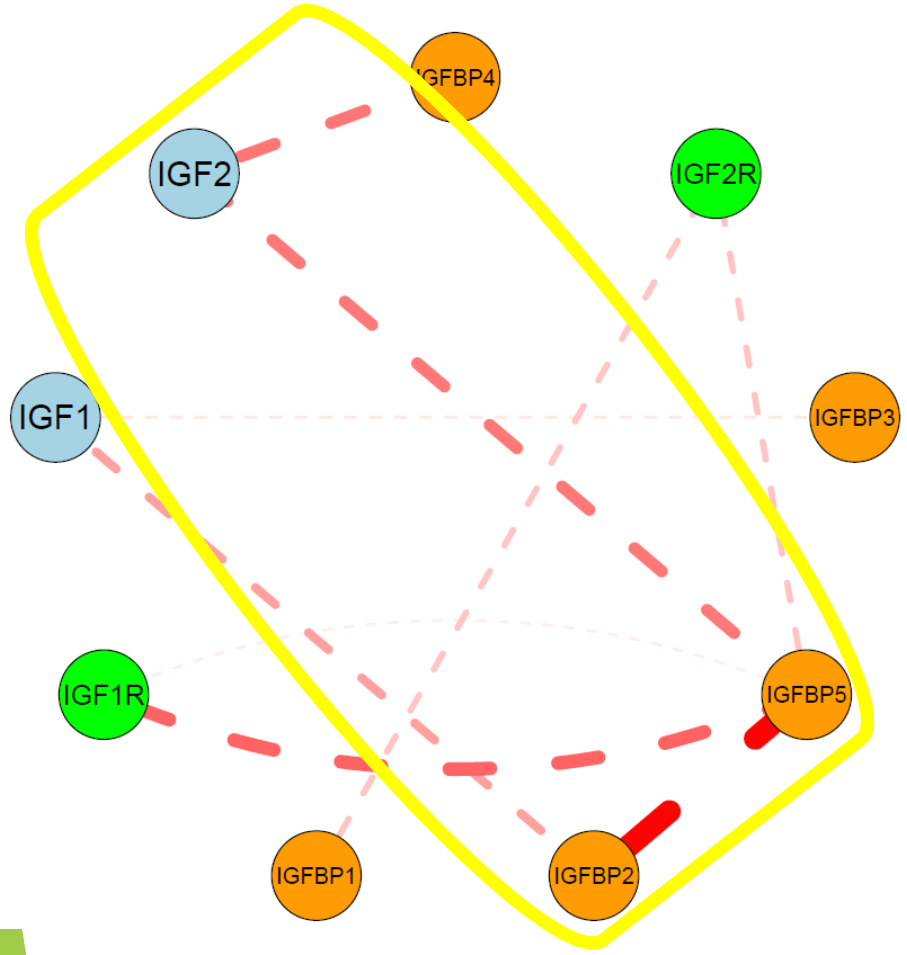


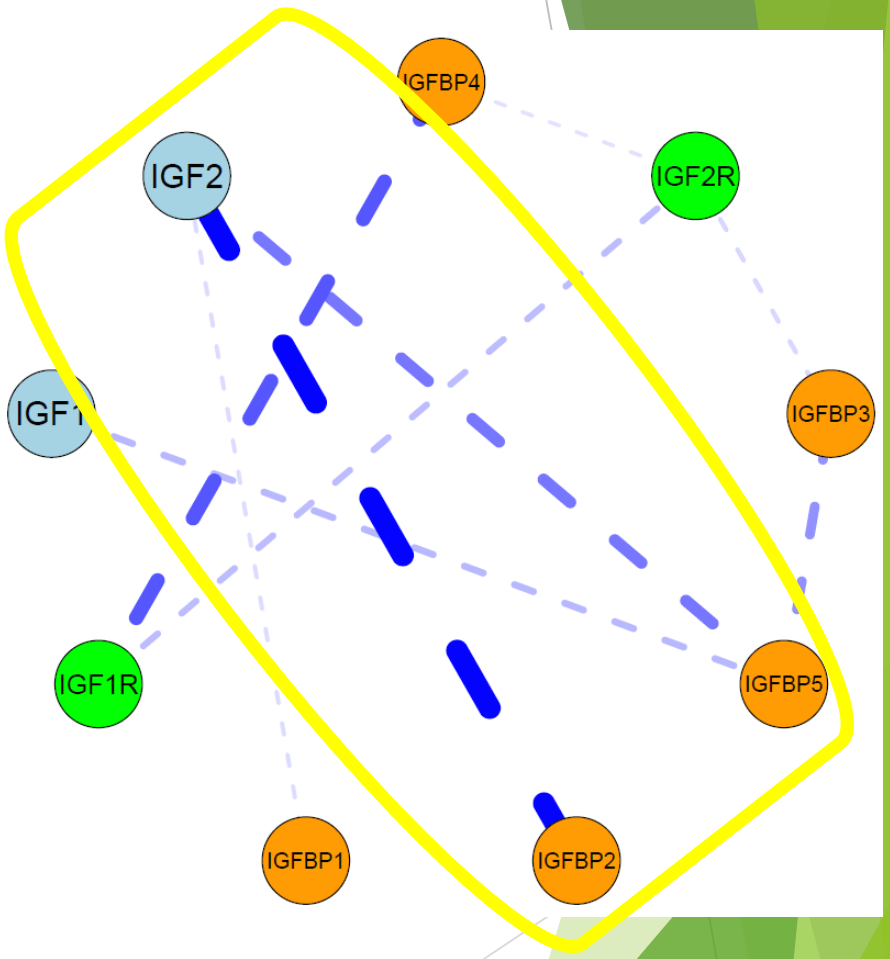
CONCLUSIONS

- Children aged over 15 years had a much more active gene-network influencing SBP compared to children aged < 15 years
- For children aged less than 15 years only males were statistically significantly associated with elevated SBP
  - IGF1-IGFBP4
  - IGF1R-IGFBP1-IGFBP3
- For females aged 15 years and over, increases to extreme SBP was observed between
  - IGF2 and IGFBP2 and IGFBP4
  - IGFBP2-IGFBP5 (and IGF1R-IGFBP5)
- For males aged 15 years and over, increases to extreme SBP was observed between
  - IGF2 and IGFBP2 and IGFBP5
  - IGF1R-IGFBP4

Females

Males





Overlapping effects between males and females for increasing elevated SBP was observed for the gene-gene interactions of IGF2-IGFBP2-IGFBP5

MEANING AND RELEVANCE

- The results attained are reasonable and align logically with current literature
- Polymorphisms within IGF2 have already been shown to influence regulation of blood pressure in obese children
- We found that this gene-network is modified with age; this in itself may be due to a number of reasons including diet, hormones and developmental growth over time, particularly post-puberty

FUTURE

- Further investigation using this method and larger datasets would be ideal to validate our findings and improve accuracy surrounding the estimates produced from these models and further improve the power to detect complex interactions

- Through characterizing the association between multiple genes and disease outcomes we will offer new insight into disease aetiology whilst providing tools for making individualized treatment decisions

ACKNOWLEDGEMENTS

Statistics in Medicine

Special Issue Paper

Received 1 November 2010, Accepted 26 July 2011 Published online 14 October 2011 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.4370

A two-stage mixed-effects model approach for gene-set analyses in candidate gene studies

Roula Tsonaka,[a*†] Annette H. M. van der Helm-van Mil[b] and Jeanine J. Houwing-Duistermaat[a]

THE UNIVERSITY OF WESTERN AUSTRALIA

Lunenfeld-Tanenbaum Research Institute

MOUNT SINAI HOSPITAL

Joseph and Wolf Lebovic Health Complex

THE UNIVERSITY of ADELAIDE

THE JOANNA BRIGGS INSTITUTE