TFS HealthScience
Clinical Research Excellence

MSR33

Improving analysis of Continuous Predictors: advantages of Fractional Polynomial transformations (FP) and interpretation of "*non-linear*" Odds Ratios (OR) or Hazard Ratios (HR)

Neus Valveny[1], Tomas O´Mahoney[2], Veronica Alfonso[3], Irene Mansilla[4], Andrew Shala[4]

1. RWE, TFS HealthScience, Barcelona, Spain; 2. RWE, TFS HealthScience, Newbridge, Ireland; 3. Biostatistics, TFS HealthScience, Barcelona, Spain; 4. Medical Writing, TFS HealthScience, Barcelona, Spain

Introduction

- In medicine, many important predictors, risk factors, confounders and/or effect-modifiers are continuous (e.g., cholesterol, blood pressure, age).
- Assessing continuous predictors with a categorical outcome (e.g., mortality) under a "linearity assumption" (same risk increase or decrease, denoted by OR or HR, per unit throughout the entire range of X values) may lead to wrong medical decisions, especially if the variable is categorized (data-driven approach).

Objective

- We aimed to review methods for analyzing continuous predictors in regression models from RWE studies, to describe advantages of FP over traditional methods, and to develop a new visualization tool to help understanding "non-linear" ORs or HRs.

Methods

- We searched in Pubmed RWE studies from the last 2 years (2020-2022) using keywords and summarized main methods, including advantages and limitations.
- We described FP methodology and developed a "FP-Risk Score Calculator" to translate model parameters into 10 risk zones (intuitive, traffic-light-like ranging from green to dark red) for the predictor values.

Results

Main methods used in RWE to test continuous predictors

- **Figure 1** shows the most common approaches: cut-points (49.7%) and untransformed variable (47.6%). FP-transformations were rarely used (0.2%).
- **Table 1** summarizes the advantages and limitations of the main approaches.
- FP was the most efficient method selecting best fit based on power and alpha error, comprising most biologically plausible risk shapes (linear/non-linear, monotonic/unimodal).

Figure 1. Main methods to test continuous predictors

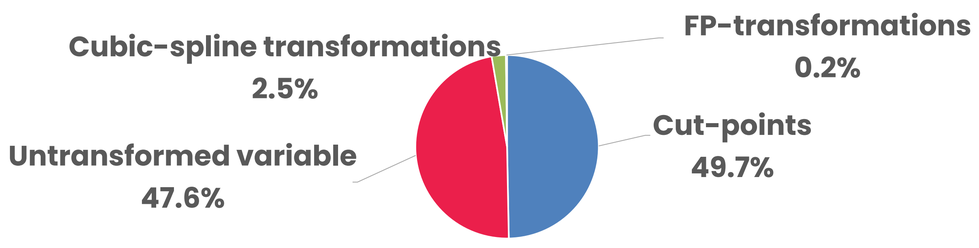


Table 1. Main methods for analyzing continuous predictors in regression models

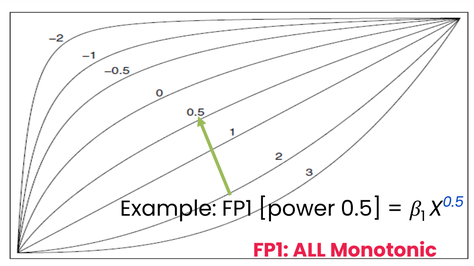
Method	Risk Shape	Main Advantages	Main Limitations
Cut-points (cut-offs from bibliography or data-driven, step risk assumption)	➢Step Risk	✓ OR/HRs easy to interpret ✓ Straightforward medical decisions/clinical algorithms	x Not biologically plausible x Grouping may lead to wrong conclusions x Loss of statistical power x Increased alpha error if multiple testing x Results may not be extrapolable to other cohorts with different "risk shapes"
Untransformed variable (continuous, linearity assumption)	➢Linear risk	✓ Linear/proportional risk is more intuitive and physicians are more familiar with it ✓ A single effect measure (OR/HR) is obtained per variable ✓ Keep statistical power	x More biologically plausible, but not entirely true ("zero" and "infinite" risks do not exist) x If non-linear/non-monotonic risk may lead to wrong conclusions x OR/HRs difficult to interpret (very small) x Results may not be extrapolable to other cohorts with different "risk shapes"
Cubic-spline transformations (curve transformations using ≤5 "cubic polynomials" with smooth merging)	➢Any shape (including multimodal)	✓ Fit all biologically plausible "risk shapes" (including most complex) ✓ Like FP for large amount of information	x OR/HRs difficult to interpret (change over X values) x Results may not be extrapolable to other cohorts with different "risk shapes" x No clear guidelines for selecting the multivariable spline-based model x More difficult to implement than FP
FP-transformations (44 single or double-term transformations using "fraction powers")	➢Monotonic or unimodal risk	✓ Keep statistical power ✓ Do not increase alpha error ✓ Fit most biologically plausible "risk shapes" ✓ Clear guidelines for multivariable models ✓ Easy to implement	x OR/HRs difficult to interpret (change over X values) x Results may not be extrapolable to other cohorts with different "risk shapes" x Cannot fit most complex shapes

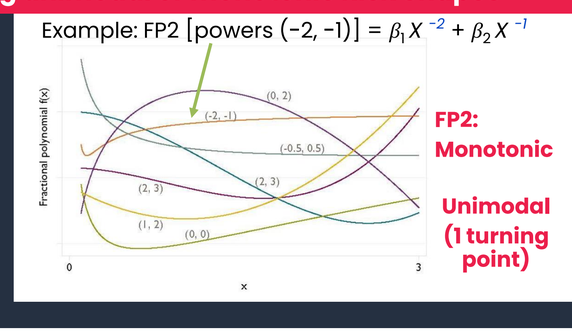
FP, fractional polynomial; HR, hazard ratio; OR, odds ratio

Fractional Polynomial transformations

- 44 transformations done:
  - ➢ 8 Single Exponents (-2, -1, -0.5, 0, 0.5, 1, 2, 3) (FP1)
  - ➢ 36 "Double" using 1 Exponent each (e.g. $\beta_1 X^{0.5} + \beta_2 X^3$) (if exponents are equal ln(x) added to second term)
- Exponents include "fraction powers" (e.g.0.5) => "**Fractional Polynomials"**
- Closed testing: Function Selection Procedure (preserves Type I Error)
- Best transformation (or linearity) is chosen
- Possible shapes reflect medical knowledge: e.g. only monotonic or unimodal forms
- Cover almost **all possible shapes**!
- Free software package for univariate or multivariate (MFP) models: https://mfp.imbi.uni-freiburg.de/
- A new R package (MFP2) will appear very soon (before end of 2023).

Figure 2. Example FP transformations assuming unimodal or monotonic Risk Shapes



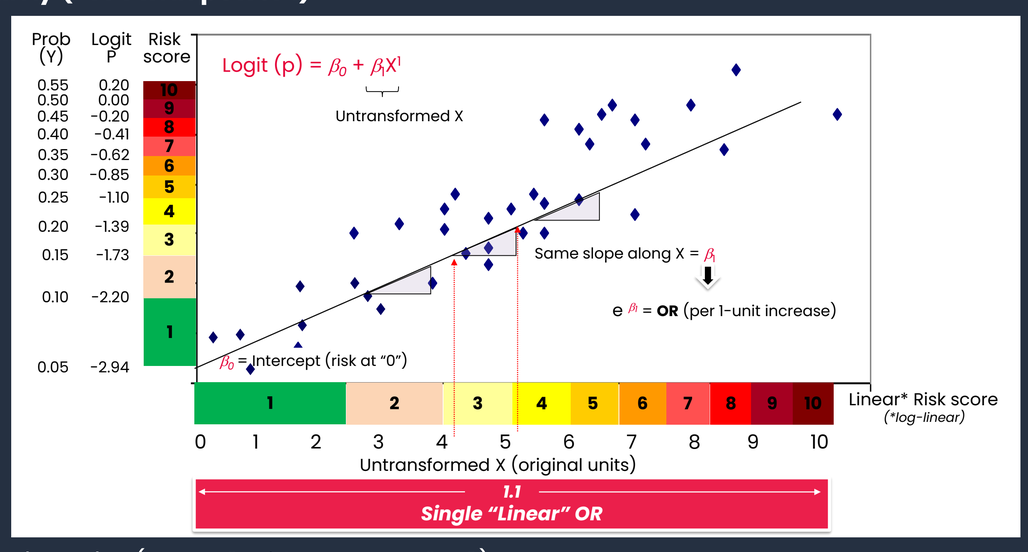


"Non-linear" Odds Ratios (OR) or Hazard Ratios (HR)

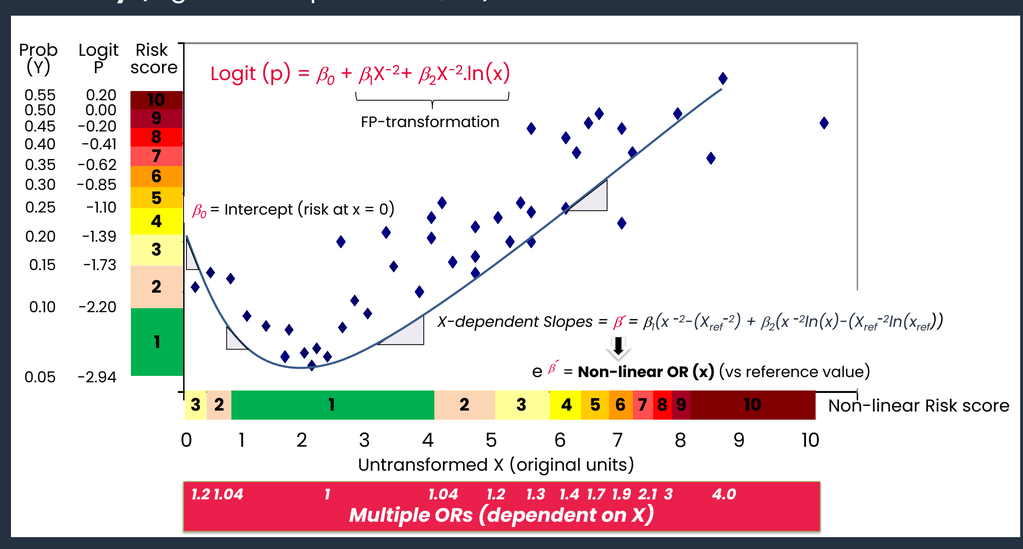
- "Non-linear" OR/HRs are x-dependent, i.e., vary along X. We developed a "FP-Risk Score Calculator" to translate model parameters into 10 Risk Zones with intuitive, traffic-light-like colours, for the observed range of X values (**Figure 3**).
- Each risk zone corresponds to relatively similar outcome probabilities. For simplicity, the OR/HR of the **mid-point** (vs reference value with lowest risk) is shown.

Figure 3. How to interpret "non-linear" OR/HR from FP transformations

A) Linearity (=FP1 with power 1)



B) Non-Linearity (e.g. FP2 with powers -2, -2)



Conclusions

- **Continuous variables should not be assumed by default to have a linear relationship with the outcome nor categorized using pre-defined or data-driven cut-points.**
- **Systematic FP-transformations are easy to implement and allow selecting the best (and simplest) "risk shape", taking into account biological plausibility (monotonic or only 1 turning point)**
- **The new FP-Risk Score Calculator divides predictor values into 10 risk zones (with 10 ORs) to facilitate clinical interpretation.**

References

Royston P, Sauerbrei W (2008) Multivariable model-building. A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous & Sonvariables (John Wiley s).
Royston,P; Ambler,G; Sauerbrei,W (1999) The use of fractional polynomials to model continuous risk variables in epidemiology. Int. J. Epidemiol.28(5): 964-974.

Acknowledgements

Prof Willi Sauerbrei and Prof Patrick Royston, who developed MFP and extensions, and initiated the STRATOS (STRengthening Analytical Thinking for Observational Studies) initiative. Research financially supported by TFS HealthScience.

Conflict of interests

All authors are employees of TFS HealthScience.

ISPOR Europe Congress. Copenhagen, 12th-15th November 2023