Local item dependence for the Rasch model

EA 4275 "Biostatistique, Pharmacoépidemiologie et Mesures Subjectives en Sante", Faculté de Pharmacie, Université de Nantes, 1 rue Gaston Veil, Nantes, France.
E-mail: Mohand-Larbi.Feddag@univ-nantes.fr

Abstract

Measurement specialist routinely assume examinee responses to test are independent of one another. However, previous research has shown that many tests contain item dependencies, and not accounting for these dependencies leads to misleading estimates of item and person parameters. In this paper, the marginal maximum likelihood estimation in Rasch model with the violation of the local independence is studied where the integrals are approximated by Gauss-Hermite quadrature. The power of the Wald test on the group effect parameter of the latent traits in cross-sectional studies is examined under the local independence and the local item dependence. The different results are illustrated with a simulation study.

Key words: Gauss-Hermite quadrature; Group effect parameter; local independence; local item dependence; Marginal maximum likelihood; Statistical power; Rasch model; Wald test.

1 Introduction

Patient Reported Outcomes (PRO) such as quality of life and other perceived health measures (pain, fatigue, stress,...) are increasingly used as important health outcomes in clinical trials or in epidemiological studies. They cannot be directly observed nor measured as other clinical or biological data and they are often collected through questionnaires with binary or polytomous items. Item response theory (IRT) models enable to model relationship between observed and latent variables where the probability of answering to each item is modelled as a function of the latent variable and item parameters. Two main statistical approaches can be used to analyze PRO data: classical test theory (CTT) based on observed scores and IRT models. The Rasch model (Rasch, 1960; Fischer and Molenaar, 1995) is the most well known used for binary responses.

Local independence of items is an assumption in Rasch model and all IRT models, that is, the item in a test should not be related to each other. Models of modern test theory imply statistical independence among responses, generally referred to as local independence. Marais and Andrich (Marais and Andrich, 2008; Christensen et al., 2013) pointed out that local independence in IRT models can be violated in two ways that are difficult to distinguish empirically and are not distinguished clearly in the literature. They distinguish between a violation of unidimensionality, which is called trait dependence, and a specific violation of statistical independence, which is called response dependence, both of which violate local independence. In this paper, we focus only on the response dependence and we call it local item dependence.

One violation of local independence occurs when the response to one item governs the response to a subsequent item. An example of local item dependence is the physical functioning subscale of the SF-36 questionnaire (Leplège et al., 2001). The items climbing one flight of stairs and climbing several flights of stairs are dependent. Similarly, the items walking one block, walking several blocks and walking more than a mile are dependent in this way. Kreiner and Christensen (2007) showed in both cases that these items are local item dependent.

Two separate statistical analyses of the data which are simulated with varying degrees of local dependence, are developed in this paper. The first uses the classical Rasch model where it is assumed that all items are statistically independent. In the second analysis, we use the model which takes into account the local item dependence. Hence, the local independence is not violated in the first analysis and violated in the second analysis.

The first aim of this paper is to investigate the effects of the violation of the assumption of local independence on the difficulty and person parameters of the Rasch model. We expect that in the data with local item dependence, the bias and the standard deviation of all the estimates will be better to those obtained under the local independence.

Statistics literature in the social, behavioral, and biomedical sciences typically stress the importance of power analysis. By definition, the power of a statistical test is the probability that its null hypothesis (H_0) will be rejected given that it is in fact false. Obviously, significance tests that lack statistical power are of limited use because they cannot reliably discriminate between the credibility of the H_0 assumption and its non rejection due to a lack of power. For cross-sectional studies comparing two groups, Hardouin et al. (2012) proposed the Raschpower procedure for the Rasch model to evaluate the power of the test of group effect. The power for detecting a prespecified group effect is determined for a given sample size, inter individual variability (variance of the latent trait) at level α (Julious, 2009; Chow, 2011).

The second aim of this paper is to evaluate the robustness of the Raschpower procedure against the violation of the local independence. Glas and Hendrawan (2005) showed that the introduction of a violation of local independence only leads to an inflation of the power in the condition where the violation is only applied for one treatment group.

The outline of the paper is as follows. Section 2 presents the dichotomous Rasch model with local item dependence. Then the marginal maximum likelihood estimation of all the parameters is illustrated by simulation study in Section 3. Section 4 is devoted to the Raschpower procedure under the two statistical analyses which are illustrated by a simulation studies. We finally conclude in Section 5.

2 The model

The Rasch model with local item dependence is defined as follows:

Let $X = (X_{ij}),\ i = 1, \ldots, N;\ j = 1, \ldots, J$ the matrix of binary variables with density distribution defined by:

$$P(X_{ij} = x_{ij} \mid \delta_j, \theta_i) = \frac{exp(x_{ij}(\theta_i - \delta_j))}{1 + exp(\theta_i - \delta_j)}, \tag{1}$$

where $\delta = (\delta_1, \ldots, \delta_J)$ is the item difficulty parameters, $\theta_1, \ldots, \theta_N$ are the latent traits supposed independent and identically distributed as a normal with mean μ and variance σ^2.

The local item dependence occurs when a person's response to an item depends on the response to a previous item. It is formalised by making a person's response on an item be a function of the person's response to a previous item. Let (j_1, j_2) a pair of items. The dependence between the response on item j_2 and the response on item j_1 is modeled by the introduction of a parameter d, $d \in \mathbb{R}$. Their conditional distribution is given by

$$P(X_{ij_2} = x_{ij_2} \mid X_{ij_1} = x_{ij_1}, \delta_{j_2}, d, \theta_i) = \frac{exp(x_{ij_2}(\theta_i - \delta_{j_2} + dx_{ij_1}))}{1 + exp(\theta_i - \delta_{j_2} + dx_{ij_1})}. \tag{2}$$

The parameter d represents the magnitude of dependence between the two items.

We note that these two items are dependents and the number of pairs of items dependent could be greater than one. Let (j_3, j_4) an another pair of items dependent with conditional distribution given by

$$P(X_{ij_4} = x_{ij_4} \mid X_{ij_3} = x_{ij_3}, \delta_{j_4}, d, \theta_i) = \frac{exp(x_{ij_4}(\theta_i - \delta_{j_4} + dx_{ij_3}))}{1 + exp(\theta_i - \delta_{j_4} + dx_{ij_3})}. \quad (3)$$

The response variables to items which are different from j_1, j_2, j_3 and j_4 are independent conditionally on the latent traits.

The model defined by Eqs (1)-(3) is the Rasch model with local item dependence. Note that if $d = 0$, all the response variables to items are conditionally independent, so the classical Rasch model holds.

This model as defined is not identifiable. So, the classical identifiability constraint we made on the parameters is $\sum_{j=1}^{J} \beta_j = 0$.

The main goal is the estimation of the difficulty parameters δ, the person parameters μ and σ, and the parameter d.

Conditionally on the latent trait θ_i, the joint probability of the variable $X_i = (X_{i1}, \ldots, X_{iJ})$ for the model defined by Eqs (1)-(3) is given as follows:

$$\begin{aligned} P(X_{i1} = x_{i1}, \ldots, X_{iJ} = x_{iJ} \mid \delta, d, \theta_i) = & \prod_{k \neq j_2, j_4} P(X_{ik} = x_{ik} \mid \delta_k, \theta_i) \\ & \times P(X_{ij_2} = x_{ij_2} \mid X_{ij_1} = x_{ij_1}, \delta_{j_2}, d, \theta_i) \\ & \times P(X_{ij_4} = x_{ij_4} \mid X_{ij_3} = x_{ij_3}, \delta_{j_4}, d, \theta_i). \end{aligned} \quad (4)$$

Hence, we deduce the marginal probability of the vector X_i as follows

$$P(X_{i1} = x_{i1}, \ldots, X_{iJ} = x_{iJ}) = \int_R P(X_{i1} = x_{i1}, \ldots, X_{iJ} = x_{iJ} \mid \delta, d, \theta_i)\ \varphi(\theta_i)\ d\theta_i, \quad (5)$$

where $\varphi(.)$ is the density function of the latent trait θ_i.

And the marginal likelihood for the defined model is given by:

$$L(\delta, \mu, \sigma, d \mid x) = \prod_{i=1}^{N} P(X_{i1} = x_{i1}, \ldots, X_{iJ} = x_{iJ}). \quad (6)$$

The integral involved in this marginal likelihood is approximated by Gauss-Hermite quadrature.

3 Simulation study

The marginal maximum likelihood estimation of the model defined above is illustrated by a simulation study where the data are simulated under the local item dependence. Then we compare the estimates of the proposed model to those obtained under the local independence. The different parameters considered in this study are fixed as follows:

- $N = 100,\ 300$ and 500
- $J = 5$ and $\delta = (-2, -1, 0, 1, 2)$
- $\mu = 0.2$, $\sigma = 1$
- pairs of dependent items : $(1, 3)$ and $(2, 4)$
- $d = 0,\ 0.2,\ 0.5$ and 1

The different estimates and their standard deviations which are based on 1000 datasets are given in the Tables 1-4, respectively for $d = 0,\ 0.2,\ 0.5$ and 1. We denote by Ind. the model with local independence and by Dep. the model with local item dependence.

Table 1 with $d = 0$, where the two models are equivalent, shows that all the estimates are unbiased for the three sample sizes considered. We note as expected that their standard deviations decrease as N increases.

For $d = 0.2$, Table 2 shows that the bias of all the estimates for the model with local independence are greater to those of the model with local item dependence. In fact, for $N = 100$, the largest bias is for δ_1 which is equal to 0.121 under the model with local independence and equal to 0.055 under the model with local item dependence. For $N = 300$, this bias is equal to 0.094 under the model with local independence and equal to 0.022 under the model with local item dependence. For both models, we note that the bias and the standard deviations are decreasing when N increases. The estimate of the parameter d is unbiased under the model with local item dependence and its standard deviation decreases as N increases. We conclude that this small value of d affect the bias of all the estimates when the local independence is violated.

With $d = 0.5$, Table 3 shows clearly that for the three values of N, the estimates of almost all the parameters are biased under the local independence and unbiased under the model which takes into account the local item dependence. We note also that the standard deviations decrease when N increases.

Table 4 shows that for $d = 1$ the estimates are strongly biased for the model with local independence and unbiased for the model with local item dependence. The estimate of the parameter d is better when N increases. We note as expected that for all the cases, the standard deviations decrease when N

Table 1
Paramer estimates and their standard deviations in parentheses for d=0.

	N=100		N=300		N=500	
parameter	Ind.	Dep.	Ind.	Dep.	Ind.	Dep.
δ_1	-2.041(.295)	-2.057(.362)	-2.016(.164)	-2.022(.203)	-2.010(.126)	-2.009(.152)
δ_2	-0.999(.239)	-1.007(.258)	-1.007(.135)	-1.010(.145)	-0.999(.101)	-0.998(.108)
δ_3	-0.004(.207)	0.001(.218)	0.001(.122)	0.003(.127)	0.001(.093)	0.000(.096)
δ_4	1.017(.223)	1.025(.246)	1.009(.130)	1.012(.143)	1.003(.095)	1.002(.104)
δ_5	2.027(.274)	2.038(.306)	2.013(.161)	2.017(.179)	2.006(.124)	2.005(.135)
μ	0.204(.152)	0.209(.169)	0.203(.088)	0.205(.098)	0.201(.067)	0.201(.073)
σ	0.991(.206)	0.999(.236)	1.004(.121)	1.007(.139)	1.002(.089)	1.002(.099)
d	*	-0.013(.498)	*	-0.006(.285)	*	0.009 (.214)

Table 2
Paramer estimates and their standard deviations in parentheses for d=0.2.

	N=100		N=300		N=500	
parameter	Ind.	Dep.	Ind.	Dep.	Ind.	Dep.
δ_1	-2.121(.308)	-2.055(.370)	-2.094(.170)	-2.022(.206)	-2.086(.131)	-2.009(.153)
δ_2	-1.041(.241)	-1.008(.263)	-1.045(.138)	-1.009(.147)	-1.038(.103)	-0.999(.110)
δ_3	0.024(.208)	0.001(.218)	0.028(.123)	0.003(.128)	0.027(.093)	0.000(.096)
δ_4	1.058(.226)	1.025(.246)	1.048(.131)	1.012(.143)	1.041(.097)	1.002(.104)
δ_5	2.080(.280)	2.038(.309)	2.063(.165)	2.017(.180)	2.055(.127)	2.005(.135)
μ	0.235(.158)	0.209(.170)	0.232(.090)	0.205(.099)	0.230(.069)	0.201(.073)
σ	1.034(.208)	0.996(.243)	1.044(.123)	1.006(.140)	1.043(.091)	1.002(.100)
d	*	0.207(.519)	*	0.201(.294)	*	0.209 (.216)

increases.

From this simulation study, we can conclude that when the parameter d increases, the bias of the parameter estimates increases under the model with local independence and remains negligible under the model with local item dependence.

Table 3
Paramer estimates and their standard deviations in parentheses for d=0.5.

	N=100		N=300		N=500	
parameter	Ind.	Dep.	Ind.	Dep.	Ind.	Dep.
δ_1	-2.223(.320)	-2.054(.377)	-2.195(.178)	-2.020(.209)	-2.188(.136)	-2.008(.157)
δ_2	-1.095(.245)	-1.012(.266)	-1.099(.142)	-1.011(.151)	-1.089(.104)	-0.999(.110)
δ_3	0.060(.210)	0.001(.219)	0.063(.125)	0.002(.127)	0.062(.094)	0.000(.096)
δ_4	1.111(.229)	1.026(.248)	1.100(.135)	1.012(.144)	1.093(.098)	1.002(.105)
δ_5	2.148(.293)	2.039(.311)	2.130(.170)	2.017(.179)	2.122(.131)	2.005(.136)
μ	0.273(.165)	0.209(.171)	0.270(.093)	0.205(.099)	0.268(.072)	0.201(.073)
σ	1.090(.215)	0.998(.247)	1.098(.127)	1.006(.141)	1.096(.094)	1.001(.101)
d	*	0.527(.671)	*	0.503(.306)	*	0.513 (.236)

Table 4
Paramer estimates and their standard deviations in parentheses for d=1.

	N=100		N=300		N=500	
parameter	Ind.	Dep.	Ind.	Dep.	Ind.	Dep.
δ_1	-2.367(.337)	-2.051(.389)	-2.343(.184)	-2.024(.216)	-2.328(.144)	-2.008(.164)
δ_2	-1.171(.251)	-1.013(.273)	-1.170(.146)	-1.009(.156)	-1.160(.107)	-0.998(.113)
δ_3	0.110(.213)	0.000(.220)	0.113(.125)	0.003(.127)	0.110(.095)	0.000(.097)
δ_4	1.185(.233)	1.025(.249)	1.174(.136)	1.013(.145)	1.164(.100)	1.002(.106)
δ_5	2.243(.304)	2.038(.314)	2.225(.177)	2.018(.180)	2.212(.135)	2.004(.136)
μ	0.327(.173)	0.209(.172)	0.324(.098)	0.205(.100)	0.320(.075)	0.201(.074)
σ	1.163(.224)	0.996(.252)	1.173(.131)	1.006(.143)	1.167(.096)	1.000(.103)
d	*	1.119(.941)	*	1.024(.350)	*	1.016 (.261)

4 Model with two groups of patients

4.1 The model

We plan to conduct a cross-sectional study for the comparison of two groups of patients. Let N_1 and N_2 be the sample sizes in each group. Let θ^1 and θ^2 be the latent traits in the first (coded 1) and in the second group (coded 2) with normal distributions respectively given by $N(-\gamma/2, \sigma^2)$ and $N(\gamma/2, \sigma^2)$. In this case, γ represents the difference between the mean values of the latent trait in the two groups.

We are interested in the comparison of the two hypotheses: $H_0: \ \gamma = 0$ vs $H_1: \ \gamma \neq 0$.

During the planning phase of the study, the values of the parameters are assumed as fixed to some hypothesis values, so it is possible to consider that the parameters δ, σ and d as known. Hence the marginal likelihood for the parameter γ is given by:

$$L(\gamma \mid \delta, \sigma, d, x) = \prod_{g=1}^{2} \prod_{i=1}^{N_g} P(X_{i1} = x_{i1}, \ldots, X_{iJ} = x_{iJ}) \tag{7}$$

where $P(X_{i1} = x_{i1}, \ldots, X_{iJ} = x_{iJ})$ is the marginal probability given by equation (5), with θ_i replaced by θ_i^1 in the first group and by θ_i^2 in the second group.

During the plannification step, the patient's responses are always unknown, so we need to determine a set of expected responses, conditionally on all the parameters fixed to their expected values. Let $X = (x^{(p)})$ the matrix of dimension $2^J \times J$, where $x^{(p)} = (x_1^{(p)}, \ldots, x_J^{(p)})$, $x_j^{(p)} = 0, 1; j = 1, \ldots, J$ is the pth binary response pattern associated to the model defined in Section 2. Let π_{pg} the probability of $x^{(p)}$ in the group g $(g = 1, 2)$, which is given by:

$$\pi_{pg} = \int_R P(X_{i1}^{(p)} = x_{i1}^{(p)}, \ldots, X_{iJ}^{(p)} = x_{iJ}^{(p)} \mid \delta, d, \theta_i^g) \ \varphi(\theta_i^g) \ d\theta_i^g. \tag{8}$$

The expected frequencies n_{pg} of each pattern p in each group g is then determined in the following way:

First we evaluate $n_{pg}^\star = floor(N_g \times \pi_{pg})$ with $floor(x) = n$ if $n \leq x < n+1$, where n is an integer. Then we calculate the number of unaffected frequencies $N_g^\star = N_g - \sum_p n_{pg}^\star$ and thereafter we compute the residual probabilities $\pi_{pg}^\star =$

$\pi_{pg} - n^{\star}_{pg}/N_g$. Then the unaffected frequencies are distributed among all the $N^{\star}_g$ patterns having the greatest values of the residual probabilities $\pi^{\star}_{pg}$, where we add 1 to the frequency. Thus $n_{pg} = n^{\star}_{pg} + 1$ for these unaffected frequencies and $n_{pg} = n^{\star}_{pg}$ for the others. Hence we construct the expected sample with size N_g, where each pattern p is repeated n_{pg} times ($p = 1, \ldots, 2^J$).

4.2 Power of the Wald test

This difference between means γ can be tested by the Wald test (see Greenland, 1983; Hardouin et al., 2012). We assume the case of typical null hypothesis that implies that there is no difference between means for the two groups. This test is performed on the two hypotheses: $H_0 : \ \gamma = 0$ and $H_1 : \ \gamma \neq 0$, and the statistic test defined by $\frac{\gamma}{\sqrt{Var(\gamma)}}$, where $Var(\gamma)$ is the variance of γ. The null hypothesis is rejected at level α if $\frac{|\hat{\gamma}|}{\sqrt{Var(\hat{\gamma})}} > z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the quantile of the cumulative standard normal distribution function, and $\hat{\gamma}$ and $Var(\hat{\gamma})$ are respectively the estimate of γ and its variance. The expected power of this test which is based on the Cramer-Rao bound is evaluated as follows:

$$1 - \hat{\beta}_{CR} = 1 - \Phi\left(z_{1-\alpha/2} - \frac{\hat{\gamma}}{\sqrt{Var(\hat{\gamma})}}\right) + \Phi\left(-z_{1-\alpha/2} - \frac{\hat{\gamma}}{\sqrt{Var(\hat{\gamma})}}\right), \quad (9)$$

where $\Phi(.)$ is the cumulative standard normal distribution function.

Assuming $\gamma > 0$, the second part of the right hand side of this equation is close to 0, thus this power is approximated as follows:

$$1 - \hat{\beta}_{CR} \simeq 1 - \Phi\left(z_{1-\alpha/2} - \frac{\gamma}{\sqrt{Var(\hat{\gamma})}}\right). \quad (10)$$

We point out that for $d = 0$, this method is the Raschpower procedure defined by Hardouin et al. (2012).

4.3 Estimation of the power using simulation

The estimation of the parameter γ using simulated data is obtained by the maximization of the marginal likelihood given by (7), where the integral is approximated by Gauss-Hermite quadrature using fifteen points. The estimate of its variance is obtained, then we deduce the significance of the Wald test under H_1 for this dataset. Hence, the power estimate denoted by $1-\beta_S$ is the rate of significant Wald tests under H_1 over the M simulated datasets. This approach is compared to the proposed one described above.

4.4 Illustration

In this section, we compare under the two models the power of the Wald test $1-\hat{\beta}_{CR}$ obtained by the proposed approach to $1-\beta_S$, the one obtained by a simulation method which is based on 1000 datasets. The parameters considered in this study are given as follows:

- $N_1 = N_2 = 50,\ 100,\ 200,\ 300$ and 500
- $J = 5$ and $\delta = (-2, -1, 0, 1, 2)$
- $\gamma = 0.2,\ 0.5$ and 0.8, $\sigma = 1$
- pairs of dependent items: $(1, 3)$ and $(2, 4)$
- $d = 0,\ 0.2,\ 0.5$ and 1

We denote by Ind. the model with local independence and by Dep. the model which takes into account the local item dependence.

We note that for all the cases considered in this study, these two powers are increasing with N_g and γ.

For $d = 0$, where the two models are equivalent, Table 5 shows that the two powers are similar. In fact, the largest difference between the two powers is equal to 0.064, which corresponds to $N_g = 300$ $(g = 1, 2)$ and $\gamma = 0.2$. With $d = 0.2$, Table 6 shows that the two powers for each model are close to each other. The power $1-\beta_S$ obtained by simulation under the model with local independence is very close to the one obtained by the model under local item dependence. In fact, the largest difference is equal to 0.007 which corresponds to $N_g = 500$ and $\gamma = 0.2$. The powers $1-\hat{\beta}_{CR}$ obtained by the two models are also close to each other. Indeed, except the case with $N_g = 500$ and $\gamma = 0.2$, where the difference between the two powers is equal to 0.101, in all other cases the difference between the powers is very small.

With $d = 0.5$, Table 7 shows that the two powers obtained under each model are comparable and very close. In fact, except the case with $N_g = 500$ and $\gamma = 0.2$ where the difference is equal to 0.139 for the Ind. model and equal to 0.078 for the Dep. model, all other cases has small difference between the two powers. The powers calculated under the Ind. model are very close to those calculated under the model with local item dependence. Table 8 with $d = 1$, shows the same results as in the previous case. The largest difference is for $1-\hat{\beta}_{CR}$ under the two models which is equal to 0.124 and corresponding to

Table 5
Power estimates for the two models with d=0.

$N_1 = N_2$	γ	$1-\beta_S$	$1-\hat{\beta}_{CR}$
50	0.2	0.090	0.096
	0.5	0.361	0.381
	0.8	0.778	0.754
100	0.2	0.146	0.149
	0.5	0.646	0.645
	0.8	0.967	0.963
200	0.2	0.300	0.289
	0.5	0.909	0.914
	0.8	1.000	1.000
300	0.2	0.368	0.304
	0.5	0.985	0.981
	0.8	1.000	1.000
500	0.2	0.564	0.592
	0.5	0.999	1.000
	0.8	1.000	1.000

$N_g = 300$ and $\gamma = 0.2$. For the other cases, the powers are very close either in each model or when we compare the two models.

When the violation is applied to two groups of patients, the power of the group effect is similar to the analogous value found in the simulations without a model violation. From this study, we can say that the violation of the local independence didn't affect the power of the Wald test for the group effect. Hence, the Raschpower procedure is robust against the violation of the local independence using two pairs of items dependent.

Table 6
Power estimates for the two models with d=0.2.

		Ind.		Dep.	
$N_1 = N_2$	γ	$1-\beta_S$	$1-\hat{\beta}_{CR}$	$1-\beta_S$	$1-\hat{\beta}_{CR}$
50	0.2	0.087	0.098	0.088	0.096
	0.5	0.362	0.382	0.362	0.380
	0.8	0.776	0.754	0.776	0.752
100	0.2	0.154	0.159	0.153	0.154
	0.5	0.651	0.652	0.650	0.648
	0.8	0.964	0.964	0.965	0.962
200	0.2	0.302	0.259	0.303	0.229
	0.5	0.908	0.913	0.909	0.908
	0.8	1.000	1.000	1.000	0.999
300	0.2	0.367	0.355	0.364	0.431
	0.5	0.983	0.983	0.982	0.980
	0.8	1.000	1.000	1.000	1.000
500	0.2	0.554	0.448	0.547	0.549
	0.5	0.999	1.000	1.000	0.999
	0.8	1.000	1.000	1.000	1.000

Table 7
Power estimates for the two models with d=0.5.

		Ind.		Dep.	
$N_1 = N_2$	γ	$1-\beta_S$	$1-\hat{\beta}_{CR}$	$1-\beta_S$	$1-\hat{\beta}_{CR}$
50	0.2	0.092	0.097	0.092	0.097
	0.5	0.364	0.379	0.360	0.378
	0.8	0.780	0.753	0.775	0.747
100	0.2	0.153	0.169	0.150	0.153
	0.5	0.645	0.652	0.643	0.642
	0.8	0.961	0.963	0.961	0.960
200	0.2	0.295	0.278	0.290	0.246
	0.5	0.904	0.914	0.900	0.908
	0.8	1.000	0.999	1.000	0.999
300	0.2	0.364	0.395	0.353	0.373
	0.5	0.983	0.983	0.982	0.981
	0.8	1.000	1.000	1.000	1.000
500	0.2	0.550	0.689	0.549	0.627
	0.5	1.000	0.999	1.000	0.999
	0.8	1.000	1.000	1.000	1.000

Table 8
Power estimates for the two models with d=1.

		Ind.		Dep.	
$N_1 = N_2$	γ	$1-\beta_S$	$1-\hat{\beta}_{CR}$	$1-\beta_S$	$1-\hat{\beta}_{CR}$
50	0.2	0.090	0.096	0.088	0.096
	0.5	0.364	0.381	0.356	0.372
	0.8	0.774	0.754	0.764	0.741
100	0.2	0.159	0.150	0.148	0.146
	0.5	0.633	0.649	0.631	0.637
	0.8	0.959	0.963	0.959	0.958
200	0.2	0.305	0.251	0.294	0.244
	0.5	0.898	0.913	0.896	0.904
	0.8	1.000	0.999	1.000	0.999
300	0.2	0.358	0.439	0.347	0.315
	0.5	0.978	0.981	0.976	0.978
	0.8	1.000	1.000	1.000	1.000
500	0.2	0.548	0.582	0.542	0.582
	0.5	1.000	0.999	0.999	0.999
	0.8	1.000	1.000	1.000	1.000

5 Conclusions

In this paper local item dependence is formalised algebraically, data are simulated with varying degrees of dependence according to this formulation, and then analysed according to the Rasch model assuming violations and no violations. The simulation study showed that the model which takes into account the local item dependence provides better estimates for all the parameters than the one wich assumes the local independence.

The Raschpower procedure was initially developped for dichotomous items in cross-sectional studies. It allowed estimating the power of Wald test to compare the difference of the means of two group of patients on a latent variable measured by a Rasch model. The approach consists in defining a planning dataset built from the probability to observe each response pattern using Gauss-Hermite quadrature. From this dataset, the difference between the means of the two groups is estimated with its variance, and the power of the test is then evaluated.

In this paper, this approach is adapted to the model which takes into account the local item dependence. Then the power of the test is evaluated under two analyses: the model with local independence and the model with local item dependence. In each analysis, this power is then compared to the one obtained by simulations.

A second remark concerns the robustness studies presented here. The simulation study showed that the power of the Wald test on the group effect remains stable when the local independence is violated compared to the one which assumes the local independence. Hence, we can conclude that the Raschpower procedure is robust against the violation of the local independence.

Above, the proposed approach was outlined in detail and investigated in simulation studies for the local item dependence to the Rasch model with binary data. However, the approach can be easily generalised in two directions. First, polytomous PRO data by using the Partial credit model defined by Master (1982). Second, the longitudinal Rasch model in longitudinal PRO binary data framework.

Acknowledgements

The authors gratefully acknowledge support from the ANR (Agence nationale de la recherche) grant N^O PRSP 00801, held at the university of Nantes.

References

Chow, S.C., Sample size calculations for clinical trials. *Wiley interdisciplinary reviews: Computational statistics*, V. 3, issue 5, 414-427 (2011).

Christensen, K.B., Kreiner, S. and Mesbah, M., Rasch models in health. *ISTE, London and John Wiley & Sons, Inc, Hoboken* (2013).

Fischer, G.H. and Molenaar, I.W., Rasch Models Foundations, recents developments and Applications. *Springer-Verlag* (1995).

Glas, C.A.W. and Hendrawan, I., Testing Linear Models for Ability Parameters in Item Response Models. *Multivariate Behavioral Research*, **40:1**, 25-51 (2005).

Greenland, S, Tests for interaction in epidemiologic studies: a review and a study of power. *Stat Med*, **2**, 243-251 (1983).

Hardouin, J.B., Amri, S., Feddag, M.L. and Sebille, V., Towards power and sample size calculations for the comparison of two groups of patients with item response theory models. *Statistics in Medicine*, **31(11-12)**, 1277-1290 (2012).

Julious, S.A., Sample sizes for clinical trials. *CRC Press- Taylor & Francis* (2009).

Kreiner, S. and Christensen, K.B., Validity and objectivity in health related scales: analysis by graphical loglinear Rasch models, in VON DAVIER M., CARSTENSEN C.H. (eds), Multivariate and Mixture Distribution RAsch Models: Extensions and Application. *Springer, New York*, pp 329-346 (2007)

Leplège, A., Ecosse, E., Pouchot, J., Coste, J. and Perneger, T., Le Questionnaire MOS SF-36. Manuel de l'utilisateur et guide d'interprétation des scores. *Editions Scientifique Techniques et Médicales* (2001).

Marais, I. and Andrich, D., Formalising dimension and response violations of local independence in the unidimensional Rasch model. *Journal of Applied Measurement*, **9**, 1-16 (2008).

Master, G., A Rasch model for partial credit scoring. *Psychometrika*, **47**, 149-174 (1982).

Rasch, G., Probabilistic models for some intelligence and attainment tests. *Danish Institute for Educational Research*. Copenhagen (1960).