

השוואת מודלים לגיבוש אבחנה ממוחשבת של לקויות למידה במערכת מת"ל

גיא עשירי-פרוסנר וענת בן-סימון

המרכז הארצי לבחינות ולהערכה

כינוס האגודה הישראלית לפסיכומטריקה, 31 ינואר 2018



רקע: מערכת מת"ל

✓ מת"ל: מערכת לאבחון תפקודי למידה

✓ מת"ל היא מערכת ממוחשבת של מבחנים ושאלונים שפותחה לצורך קביעת זכאות להתאמות בבחינות למועמדים ולומדים במוסדות להשכלה גבוהה.

✓ המערכת נועדה לאבחן את הלקויות : דיסקסיה, דיסקלקוליה, דיסגרפיה והפרעת קשב וריכוז.

✓ מת"ל כוללת שני שאלונים ו-20 מבחנים בתחומים הבאים: שפה, תפקודים כמותיים, קשב, זיכרון, תפיסה ומהירות עיבוד כללית.

רקע: סוגי האבחנה במת"ל

✓ לכל לקות מופקים שני סוגי אבחנה:

- אבחנה ממוחשבת - מבוססת על הביצוע בכלי האבחון של מת"ל בלבד.
 - אבחנה קלינית - נעשית על ידי מאבחנת מומחית ונשענת על נתונים ממגוון רחב של מקורות מידע ובכלל זה הביצוע במבחני מת"ל.
- ✓ עבור כל לקות, האבחנה (קיימת לקות \ ללא לקות) היא משתנה בינארי, לכן נעשה שימוש ברגרסיה לוגיסטית.

רקע: רגרסיה לוגיסטית

מודל סטטיסטי המשמש לחיזוי ערכו של משתנה בינארי y_i על בסיס ערכיהם של משתנים רציפים x_{1i}, \dots, x_{pi} , לדוגמא:

	x1	x2	x3	x4	x5	y_dyslexia
1	-2.05	-4.02	-3.52	-1.30	-0.89	0
2	-2.24	-2.79	-1.30	-0.63	-0.22	1
3	-1.75	-1.01	-3.35	-1.48	-0.79	0
4	-1.32	-2.54	-1.20	-1.50	-0.97	0
5	-0.73	0.24	-0.77	-1.60	-1.25	1
6	-3.08	-3.17	-3.56	-1.55	-3.01	1

לכל תצפית i מחושבים האומד $\theta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}$ והסתברות להימצאות לקות באופן

$$P(y_i = 1|x_i) = \pi_i = \frac{e^{\theta_i}}{1 + e^{\theta_i}}$$

התחזית לכל תצפית נקבעת לפי נקודת החתך c , באופן $\hat{y}_i = \begin{cases} 1 & \hat{\pi}_i \geq c \\ 0 & \hat{\pi}_i < c \end{cases}$

רקע: פיתוח מודל רגרסיה לוגיסטית

שלבי פיתוח

1. מציאת אומד לוקטור המקדמים β ע"י פתרון המודל על מדגם אימון
2. בחינת דיוק המודל ע"י הרצתו עבור אוסף נתונים (מדגם תיקוף);
בהתבסס על מדגם התיקוף - מציאת נקודת החתך האופטימלית c
לסיווג פונים ללקויים/לא לקויים
3. יישום המודל עבור פונים חדשים

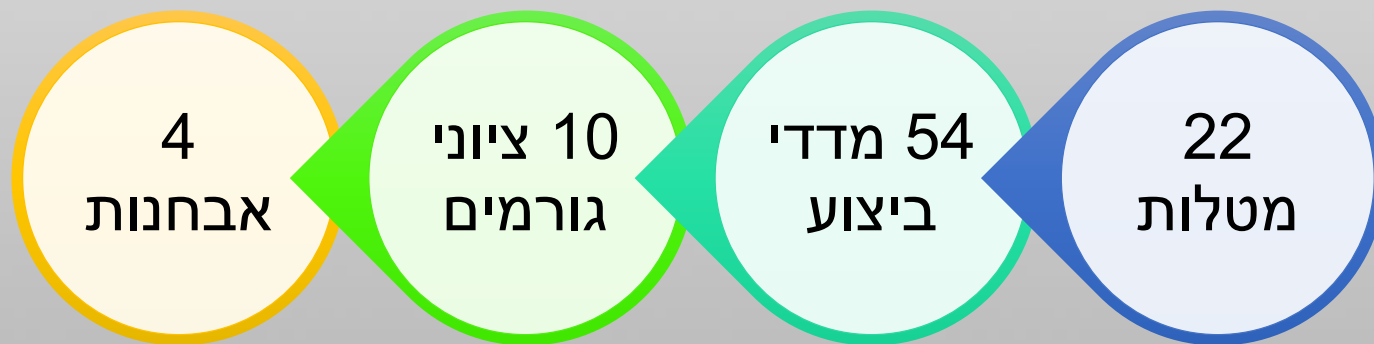
רקע: תהליך הפקת הציונים במת"ל

✓ **מדדי ביצוע (גלם):**

(למשל, דיוק וזמן תגובה)

✓ **ציוני גורמים: 10** ציונים המחושבים מתוך מדדי הביצוע המתוקננים בהתבסס על ניתוח גורמים.

✓ **אבחנות ממוחשבות: 4** מודלים של רגרסיה לוגיסטית (אחד לכל לקות) המבוססים על ציוני הגורמים.



• רגרסיה
לוגיסטית

• ניתוח
גורמים

• הפקת ציוני
תקן ביחס
לנורמות

מטרת המחקר

✓ השוואת דיוק הניבוי של שלושה מודלים של רגרסיה לוגיסטית, הנבדלים זה מזה במדגמים המשמשים להפקת מודל הניבוי.

מדגם

המחקר בוצע תוך שימוש בצירופים שונים של שני בסיסי נתונים:

✓ מדגם הנורמות - 508 משתתפים ללא לקויות למידה או הפרעת קשב וריכוז

✓ מדגם מחקר התיקוף החוזר - 563 משתתפים שפנו לאבחון מבוסס מת"ל בין 2008 ל-2011

לכל לקות הוגדרו שלוש קבוצות משתתפים:

1. משתתפי מדגם הנורמות (ללא לקות)
2. משתתפי מדגם מחקר התיקוף החוזר (ללא הלקות שבמוקד הבדיקה)
3. הקבוצה הקלינית - משתתפים ממדגם התיקוף החוזר שאובחנו כבעלי הלקות שבמוקד הניבוי: דיסלקסיה ($n=285$), דיסגרפיה ($n=255$), דיסקלקוליה ($n=131$) והפרעת קשב וריכוז ($n=313$)

שיטה: גישות

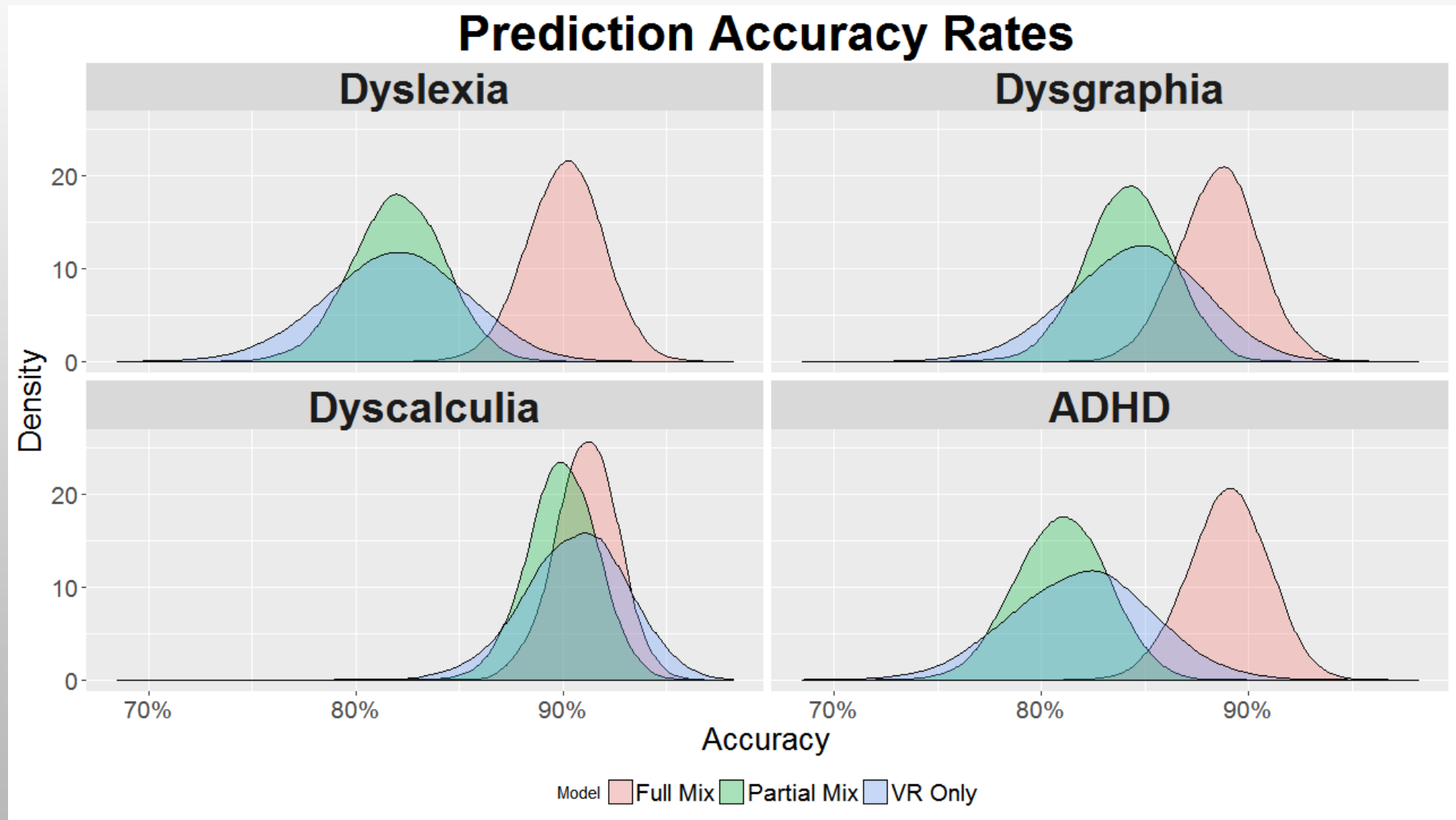
נבחנו שלוש גישות לבניית מדגמי האימון והתיקוף:

- ✓ **Full Mix** - הן קבוצת האימון והן קבוצת התיקוף מבוססות על שילוב של שני המדגמים (מדגם הנורמות ומדגם מחקר תיקוף חוזר).
- ✓ **Partial Mix** - קבוצת האימון מבוססת על שילוב של שני המדגמים ואילו קבוצת התיקוף מבוססת על מדגם מחקר התיקוף החוזר בלבד.
- ✓ **VR Only** - שתי הקבוצות מבוססות על מדגם מחקר התיקוף החוזר בלבד.

שיטה: הליך ההשוואה

- ✓ שלוש הגישות השונות נבחנו עבור כל לקות.
- ✓ עבור כל גישה, בוצעה חלוקה אקראית של הנתונים לשתי קבוצות זרות ביחס 4:1 (80% קב' אימון, 20% קב' תיקוף), תוך שמירה על יחס קבוע בין משתתפים עם וללא לקות בשתי הקבוצות.
- ✓ לכל צירוף של לקות וגישה בוצעו 10,000 חזרות (איטרציות) על חלוקת המדגם לקבוצות אימון ותיקוף.
- ✓ ההשוואה בין הגישות השונות בדיוק הניבוי של המודל נעשתה תוך שימוש באבחנה הקלינית כקריטריון ותוך התייחסות לשני סוגי הטעות (true negative, false positive).

התפלגות אחוז דיוק הניבוי של שלוש הגישות לפי לקות



תוצאות: אחוז דיוק ניבוי ממוצע לפי גישה ולקות

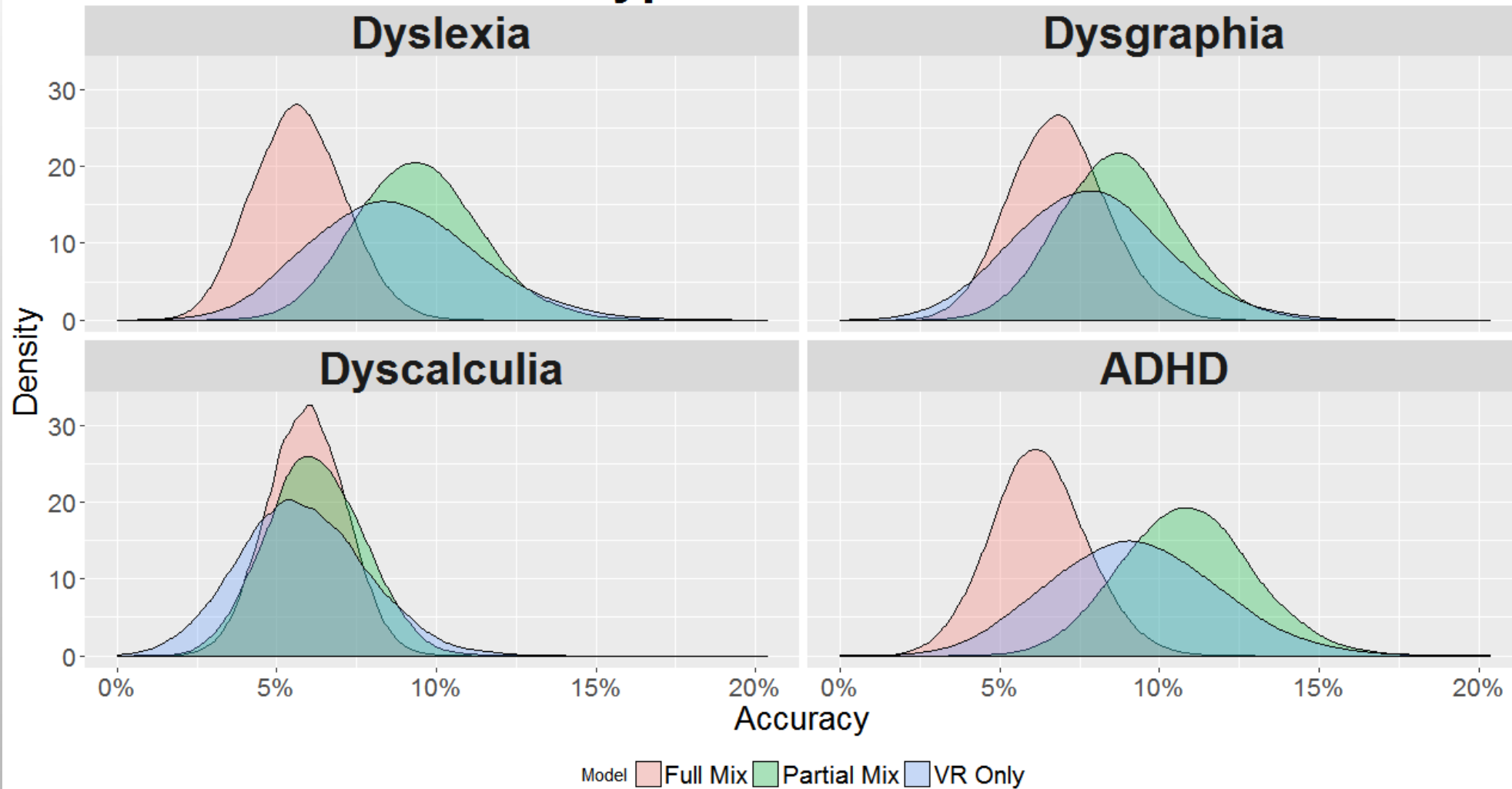
VR Only	Partial Mix	Full Mix	
80.6	81.9	87.9	דיסלקסיה
83.2	83.6	86.7	דיסגרפיה
89.2	89.8	89.1	דיסקלקוליה
80.7	80.7	87.0	הפרעת קשב וריכוז

באופן כללי ממוצע הדיוק הגבוה ביותר התקבל עבור גישת Full Mix. ✓

ממוצע הטעויות מסוג שני (true negative כלומר פונה עם לקות שזוהה כלא-לקוי) היה הנמוך ביותר עבור גישת Full Mix. ✓

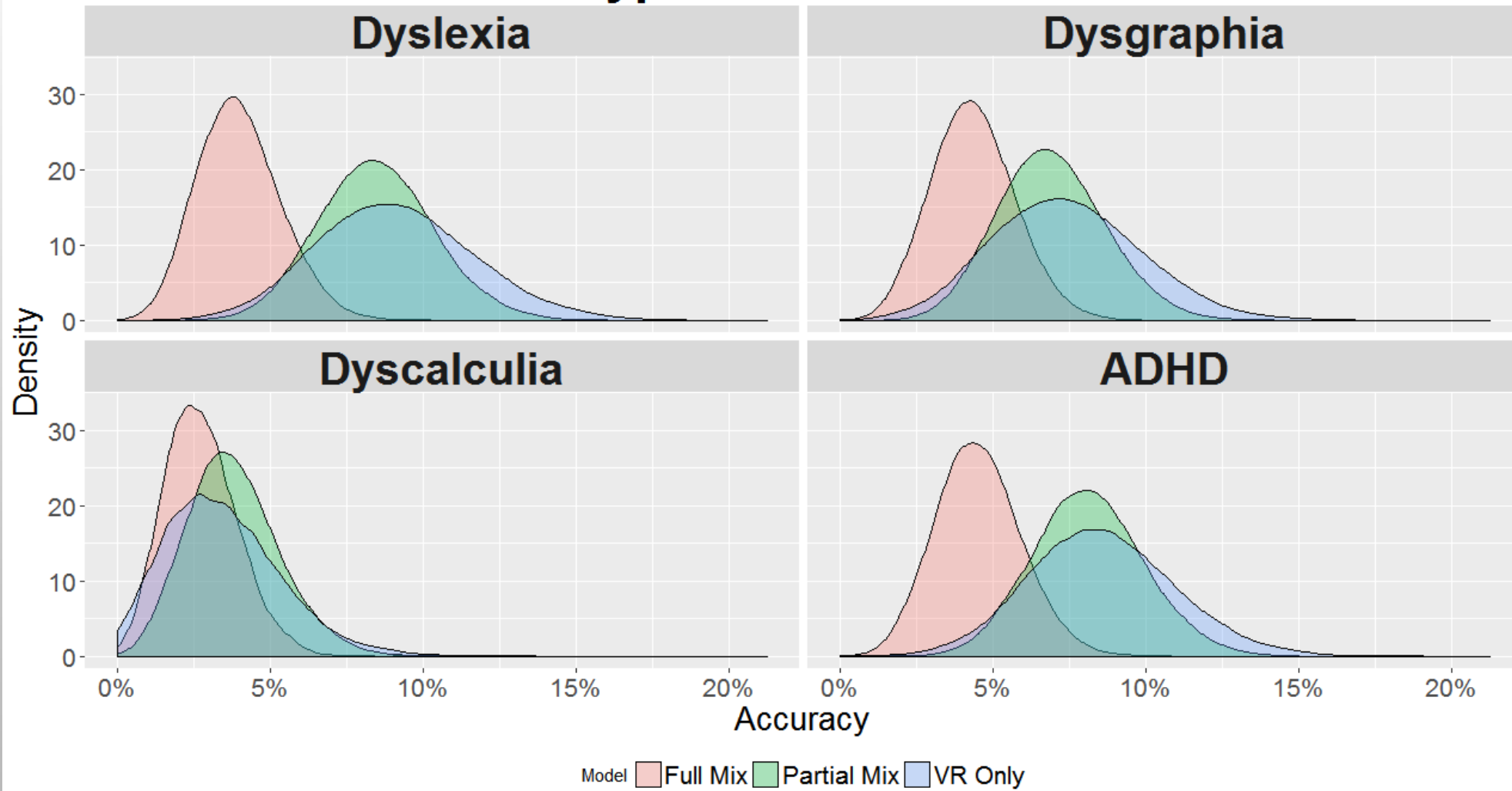
התפלגות טעות מסוג ראשון (false positive)

Type I Error Rates



התפלגות טעויות מסוג שני (true negative)

Type II Error Rates



תיקוף המודל מול מדגם נוסף

מודלים: (1) מודל אופטימלי בגישת Full Mix;
(2) מודל אופטימלי בגישת VR Only

מדגם: 1,841 פונים לאבחון מבוסס מת"ל שלא נכללו בשני המדגמים
הקודמים

קריטריון: אבחנה ממוחשבת

אחוז טעות מסוג שני		אחוז דיוק לפי מודל		
VR Only	Full Mix	VR Only	Full Mix	
6.1	5.2	87.3	87.5	דיסלקסיה
8.5	8.6	81.3	82.9	דיסגרפיה
1.3	1.7	88.8	90.3	דיסקלקוליה
9.6	8.9	86.0	88.1	הפרעת קשב וריכוז

סיכום

- ✓ מודלי הניבוי בגישת Full Mix, בהם נעשה שימוש הן במדגם הנורמות (משתתפים ללא לקות) והן במדגם התיקוף החוזר, השיגו בממוצע את אחוזי הדיוק הגבוהים ביותר בניבוי ארבע הלקויות המאובחנות על ידי מת"ל.
- ✓ מודל בגישה זו השיג תוצאות טובות יותר גם כאשר יושם על המדגם המורחב.
- ✓ הסברים אפשריים לממצאים:
 - הגדלת מספר התצפיות בקבוצת האימון מניבה מודל מדויק יותר.
 - תצפיות מדגם הנורמות מרחיבות את התחום ממנו מגיעים ערכי ציוני הגורמים (המנבאים).

תודה על
ההקשבה,
שאלות



guy@nite.org.il

מאלו