



30.4.2020

### מודל הערכת הדבקת ילדים המבוסס על בתי אב בבני ברק

#### רקע:

אחד המאפיינים הבולטים של מגיפת COVID-19 הוא אחוז נמוך של ילדים חיוביים מכלל חולי קורונה, ונמוך מחלקם היחסי באוכלוסייה. ישנן מספר סיבות אפשריות לכך:

- ילדים פחות רגישים ממבוגרים ולכן נדבקים פחות.
- ילדים שנדבקו מפתחים פחות תסמינים ולכן הם פחות מאובחנים.
- דפוסי הדבקה דיפרנציאליים (assortative mixing) וסגירת בתי הספר המפחיתה את כמות המגעים בין הילדים (Davies et al. 2020).

#### תמצית סקירת ספרות:

מסקירת הספרות שערכנו ביחס לסיכויי הידבקות של ילדים ב-COVID-19, עולה כי אחוז הילדים המאובחנים נמוך ביחס לחלקם באוכלוסייה. מתוך כלל המאובחנים אחוז הילדים בסין עומד על 0.9%, בארה"ב על 5% ובאיטליה על 1.2% (Lee et al.). החוקרים משערים כי האחוז הנמוך של הילדים שאובחנו נובע מכך שהם אינם סימפטומטיים או בעלי סימפטומים קלים ולכן הסיכוי שיופנו לבדיקה ויאובחנו נמוך יותר בהשוואה למבוגרים. חלק משמעותי מהילדים הם א-סימפטומטיים או בעלי סימפטומים קלים (Torsdahl et al., 2020). אחוז הא-סימפטומטיים המדווח נע בין 13-80% (Davies et al.; Dong et al.), ובילדים בהם נמצאו סימפטומים הם היו קלים יותר מאשר בקרב מבוגרים (Lee et al., 2020; Ludvigsson, 2020).

בסקירת הספרות לגבי סיכויי ילד להידבק, נמצא כי חלק מהחוקרים סוברים כי הסיכוי של ילד להידבק הוא כמחצית מזה של מבוגר (Gudbjartsson et al., 2020; Zhang et al., 2020) בעוד שאחרים מעריכים שהסיכוי של ילד להידבק דומה לזה של מבוגר (Bi et al., 2020). במעט מקרים (הולנד ואוסטרליה) פורסם תיעוד לכך שילדים חולים לא הדביקו אנשים שהיו איתם במגע קרוב. עם זאת, לא נמצא מחקר כמותי סדור של סיכויי ההדבקה על ידי ילדים חולים. חלק מהמחקרים מעלים השערה שהעדר סימפטומים ואבחון חלקי של הילדים החולים מביאים לכך שהילדים מהווים חוליה משמעותית בשרשראות ההדבקה (Kelvin et al., 2020). במסגרת עבודה זו אנו מתמקדים בניסיון לכמת את תפקיד הילדים בהפצת המחלה. בשל העובדה שלא ניתן לבצע ניסויים מבוקרים של בחינת שרשרת ההדבקה פרוספקטיבית, הכלי העיקרי בכדי לאמוד את אפקט ההדבקה, הוא שימוש במודלים מתמטיים וסטטיסטיים. נדגיש שבעבודה זו איננו מתייחסים כלל למשמעות הקלינית של התחלואה באוכלוסייה זו. (סקירת ספרות בנספח 6).

#### מטרות:

השאלה עליה התבקשנו לענות היא האם מאפייני ההדבקה של ילדים שונים ממבוגרים. בפרט, ניסינו לענות על שתי השאלות הבאות:

1. רגישות להידבקות: האם ילדים נדבקים (ממבוגרים או מילדים) יותר או פחות ממבוגרים?
2. כושר הדבקה: האם ילדים מדביקים אחרים (מבוגרים או ילדים) יותר או פחות ממבוגרים?



### שיטות:

בחרנו בשתי שיטות על מנת לענות על שאלות המחקר. השיטה הראשונה פשוטה ליישום ומבוססת על מודל שמניח שאין הדבקות שניוניות במשפחה (מודל בינומי). השיטה השנייה מבוססת על מודל המתאר את תהליך התפשטות המגפה בתוך משפחה, אין בה הנחות לגבי הדבקות שניוניות, והיא יותר מורכבת מתמטית וחישובית. פירוט המודלים מופיע בהמשך המסמך. פרטים טכניים ניתן למצוא בנספחים 1-3.

על מנת להפעיל שיטות אלו נדרשנו תחילה להגדיר את חתך הגיל על פיו יקבע האם פרט שנבדק הוא ילד או מבוגר. על מנת לבדוק רגישויות של התוצאות להגדרת חתך הגיל הקובעת מיהו ילד ומיהו מבוגר, בדקנו מספר אפשרויות חיתוך קבוצת הגיל, ולבסוף נבחרו שני חתכים להגדרת קבוצת הילדים: קבוצת ילדים בגיל 0-9, וקבוצת ילדים בגיל 0-19. הניתוחים בדוח זה מבוצעים עבור שני חיתוכים אלו. לאחר שהוגדרו הקבוצות, בחנו את המשפחות בהן החולה הראשון הוא ילד, וכאלה בהן הוא מבוגר. הנתונים מתבססים על שני מקורות מידע: מאגר נתונים אודות כלל תושבי בני ברק, ומאגר נתוני קורונה של משרד הבריאות. פירוט לגבי הנתונים בהמשך המסמך.

### המודל:

בבואנו לדון בהטרוגניות של אוכלוסייה מבחינת העברת מחלה מדבקת, יש להבחין בין שני ממדים: רגישות להידבקות (susceptibility) וכושר ההדבקה (infectivity). פרט א' ייחשב לרגיש יותר מפרט ב' אם ההסתברות להידבקות של פרט א' גבוהה מהסתברות ההידבקות של פרט ב', בחשיפה לאותו גורם הדבקה, כאשר היחס בין רמות הרגישות של שני הפרטים מוגדר כיחס בין ההסתברויות הללו. פרט א' ייחשב כבעל כושר הדבקה גבוה יותר מפרט ב' אם, במפגש שלהם עם פרטים המועדים להדבקה ובעלי רמת רגישות זהה, ההסתברות של פרט א' להדביק גבוהה יותר מההסתברות של פרט ב' להדביק, והיחס בין כושר ההדבקה של שני הפרטים מוגדר כיחס בין ההסתברויות הללו.

על מנת להעריך הבדלים בין רגישות להידבקות וכושר הדבקה בקבוצות גיל שונות, בהעדר אפשרות לערוך ניסויים מבוקרים, עלינו להשתמש בנתונים אמפיריים ובמודלים מתמטיים. באמצעות מודלים אלו ניתן לבחון את ההשלכות של הנחות שונות לגבי הבדלים ברגישות להדבקה וכושר ההדבקה על מספר הנדבקים הצפוי, ולמצוא את הפרמטרים שעבורם מתקבלת התאמה מיטבית לנתונים האמפיריים. במקרה שלנו הנתונים האמפיריים הם לגבי הדבקה בתוך משקי בית הכוללים: גילאים של בני הבית ונתוני בדיקות שנערכו לבני הבית, אך יש בעייה של אמינות לגבי הנתון מיהו הנדבק הראשון במשפחה (בנתונים אלו נעשה שימוש בהתאמת המודל) ויש חוסר ודאות לגבי מי הדביק את מי במשפחה (התאמת המודל אינה עושה שימוש בנתונים מסוג זה).

יש לציין שכל מודל מתמטי נסמך על הנחות, ולכן המסקנות המתקבלות ממודל נשענות לא רק על נתונים אמפיריים אלא גם על הנחות אלו. בחנו שני מודלים על מנת לבדוק באיזו מידה הנחות שונות יובילו למסקנות שונות לגבי הבדלים ברגישות להידבקות ובכושר הדבקה כתלות בגיל.

במודל הראשון, להלן המודל הבינומי, הנחנו שכל בני הבית שנדבקו, הודבקו על-ידי "מקרה האינדקס" (index case) - המדביק הראשון במשק הבית. במילים אחרות, מודל זה מניח שאין הדבקות שניוניות. אפשרנו רמות הדבקה שונות בהתאם לקבוצת הגיל (מבוגר המדביק ילד, ילד המדביק מבוגר, מבוגר המדביק מבוגר, וילד המדביק ילד). תחת מודל פשוט זה קל להעריך את פרמטרי ההדבקה (ראה נספח 1).



המודל השני, להלן מודל ההדבקות במשפחה, מסיר את ההנחה שכל בני הבית נדבקו מ"מקרה האינדקס", ומתאר תהליך סטוכסטי של הדבקה בין בני הבית השונים, בהתאם לפרמטרי רגישות להידבקות וכושר הדבקה של שתי קבוצות הגיל. פרמטרי הזמן הקשורים לתהליך ההדבקה (generation time distribution) מוערכים על סמך נתונים שהתקבלו מחקירות אפידמיולוגיות ומהספרות המדעית. על מנת להעריך באופן שיטתי את הפרמטרים המתאימים בעזרת הנתונים שברשותנו, נבנתה פונקציית הנראות (likelihood) של המודל, המתארת את המידה שבה התנהגות המודל תואמת את הנתונים האמפיריים, כתלות בפרמטרים. חישוב פונקציית הנראות נסמך על הרצה של מאות אלפי סימולציות של המודל עבור משפחות עם הרכבים שונים, ובעזרת חישוב זה אנו מעריכים את הפרמטרים התואמים את הנתונים באופן המיטבי.

### הנתונים ששימשו באנליזה :

הנתונים נלקחו משני מקורות מידע. מקור הנתונים הראשון הינו מאגר מידע המכיל את מספרי הזהות וכתובות המגורים של כל תושבי העיר בני ברק אשר נולדו עד תאריך 24.03.2020. התושבים חולקו למשפחות עפ"י כתובת מגוריהם ושם משפחתם. מקור הנתונים השני הינו מאגר הנתונים של משרד הבריאות הכולל תוצאות בדיקות המעבדה לזיהוי קורונה, והחקירות האפידמיולוגיות המתבצעות ע"י המשרד.

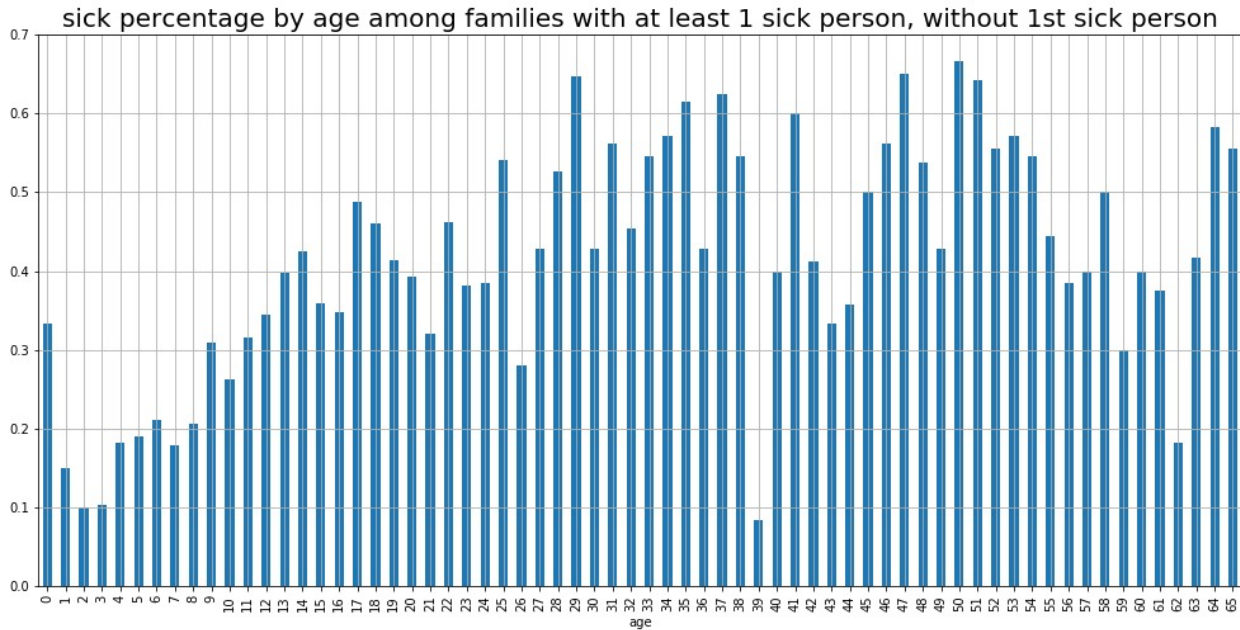
התבצעה הצלבה בין שני המאגרים עפ"י מספרי הזהות ונשלפו כל האנשים אשר עברו בדיקת קורונה. משפחות שבהן כל בני המשפחה עברו לפחות בדיקת קורונה אחת עד לתאריך 23.04.2020, ונמצא בהן לפחות בן משפחה אחד חיובי נכנסו לעיבוד הנתונים. בסה"כ הכיל הקובץ ששימש לעיבוד הנתונים נכון למועד השליפה, 746 משפחות, כולל משפחות עם אדם אחד שמתגורר בגפו. נתוני הבדיקות והחקירות מעודכנים לתאריך 25.04.2020. נכון להיום (30.4), יש עוד כמאה משפחות נוספות שניתן להוסיף למדגם, וכן התעדכנו נתונים נוספים על 746 המשפחות המקוריות (כמו בדיקות נוספות שנעשו בזמן הזה). היות והמאגר מתעדכן באופן שוטף, נתונים אלו לא התווספו לדוח זה, וישולבו בעיבוד נתונים נוסף שיתבצע בימים הקרובים על הנתונים המעודכנים ביותר שיהיו קיימים במועד הדוח הבא.

### תוצאות:

גודל בתי האב במשפחות שנדגמו במלואן עם חולה חיובי אחד לפחות נעו בין 1-15 בני משפחה לכל בית אב. בסופו של דבר לאחר שהורדנו את בתי האב שכללו אדם אחד בלבד וכאלו שהיה חסר להם גיל, נותרו 562 משפחות עם 2,823 אנשים. אוכלוסיה זו כללה 25.7% בקבוצת הגיל 0-9, 26.6% בקבוצת הגיל 10-19 ו 47.7% בקבוצת הגיל 20+. אחוז החולים האסימפטומטיים או פרה-סימפטומטיים במדגם זה בקרב מבוגרים עומד על 12.9%, בילדים עד גיל 20 עומד על 30.1% ובילדים עד גיל 10 עומד על 31.8%. בבדיקה שערכנו לגבי סיכויי התחלואה בבתי אב, כתלות בגיל ובהינתן בן משפחה חיובי, ולאחר שהורדנו את החולה הראשון בכל בית אב, ניתן לראות בגרף שלהלן (איור 1), שילדים נוטים לחלות פחות, אך הסיכוי שלהם לחלות, עולה עם הגיל. יחד עם זאת, ניתן לראות שתינוקות עד גיל שנה, הם בעלי סיכוי גבוה לחלות ביחס לשאר הילדים במשפחה. הסבר אפשרי הוא שתינוקות דורשים טיפול צמוד/הנקה מהוריהם, ולכן אם ההורים חולים (ברוב המשפחות במדגם), הסיכוי של התינוקות לחלות גדול בהשוואה לשאר הילדים.

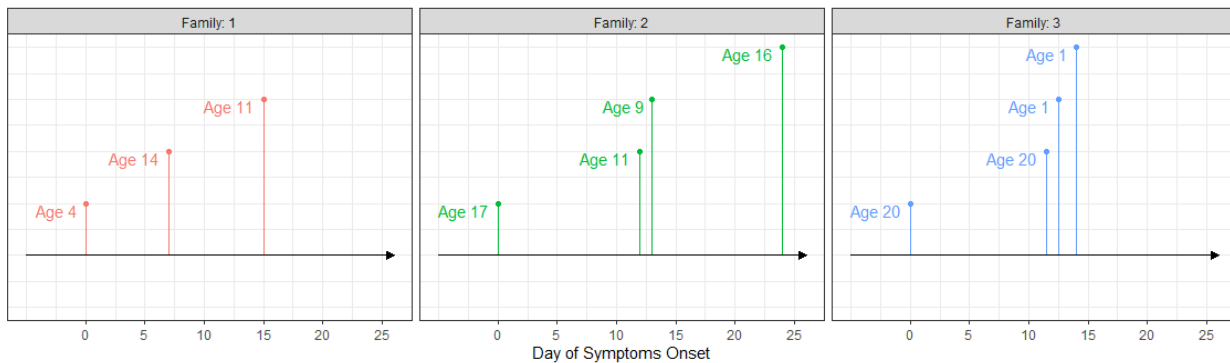


## איור 1



בדיקה של בתי אב בהם היו רק ילדים חולים (ללא חולים מבוגרים, סה"כ 66 משפחות בישראל, 12 במדגם), ממחישה כי יש אפשרות שילדים מדבקים, ובנוסף, טווח הזמנים הגדול בין הנדבק הראשון לאחרון עשוי להצביע על כך שקיימות הדבקות שניוניות. להלן 3 גרפים (איור 2) המדגימים את התופעה המתוארת לעיל, בשלוש משפחות שונות.

## איור 2





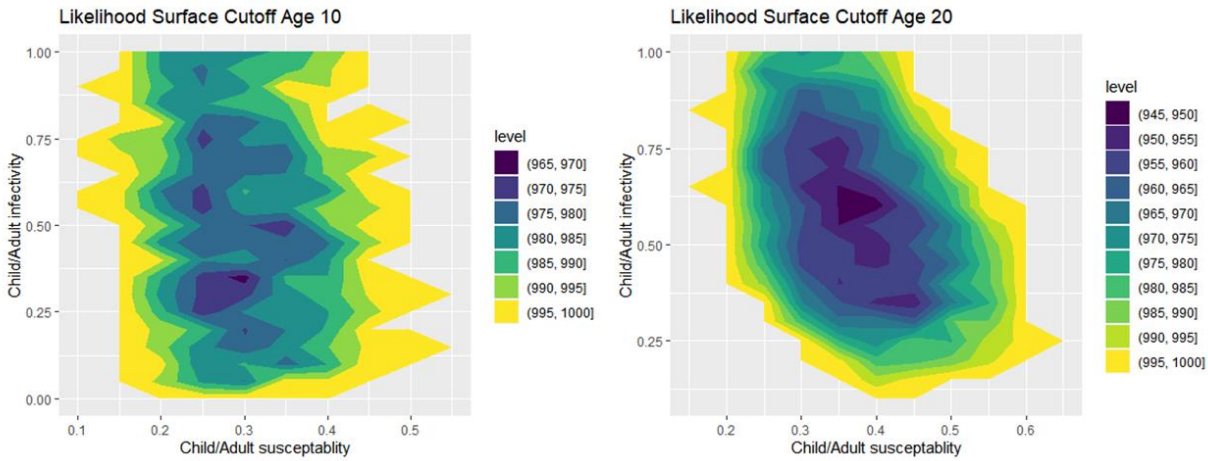
**תוצאות ניתוח הנתונים:**

מתוצאות המודל הבינומי, שמניח הדבקה אך ורק על ידי מקרה האינדקס, מתקבל כי הסיכוי של ילד להידבק קטן מהסיכוי של מבוגר, כאשר ההבדל בולט יותר עבור ילדים עד גיל 10. לעומת זאת, קשה להכריע האם ילדים מדביקים יותר או פחות ממבוגרים בשל רווחי סמך גדולים (ראה נספח 1).

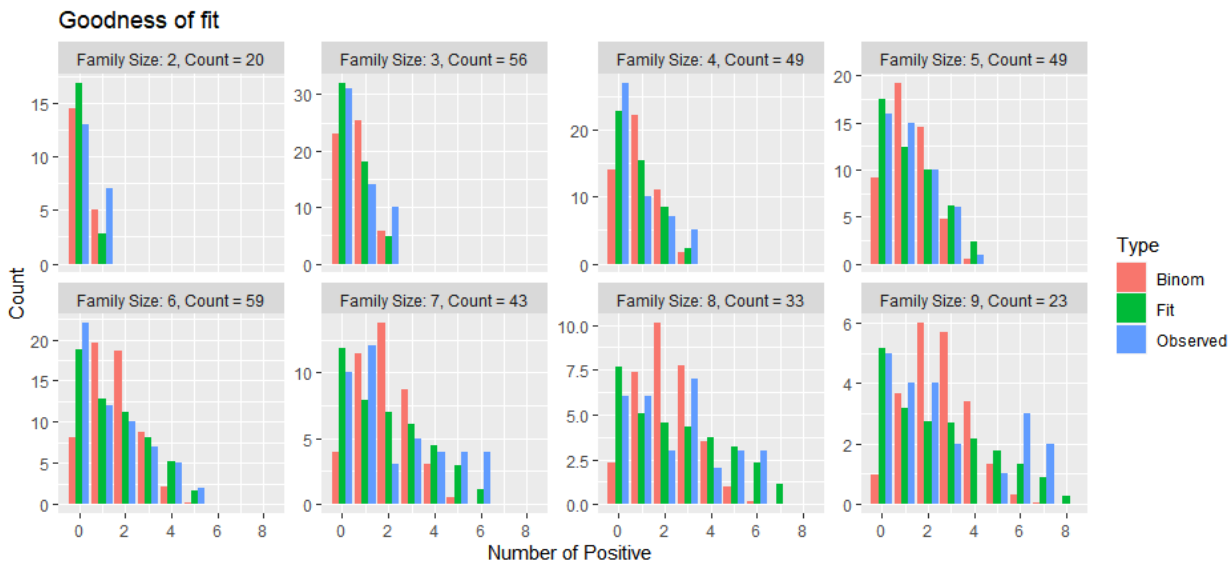
גם במודל ההדבקות במשפחה מתקבל כי הסיכוי של ילדים להידבק קטן מהסיכוי של מבוגרים, כאשר לילדים עד גיל 20 הסיכוי להידבק הוא 30-50% ביחס למבוגרים, וילדים עד גיל 10 הסיכוי להידבק הוא 20-40% ביחס למבוגרים. היכולת של ילדים להדביק בשני החתוכים, קשה יותר לאמידה ונעה בין 25% ל-75% ביחס ליכולת של מבוגרים (ראה איור 3).

על אף שבמודל ההדבקות במשפחה ישנם שלושה פרמטרים ובמודל הבינומי ישנם ארבעה פרמטרים (יותר דרגות חופש), מודל ההדבקות במשפחה תואם טוב יותר את הנתונים (ראה איור 4).

**איור 3**



**איור 4**





### מגבלות העבודה:

- הניתוח מבוסס על מדגם מתוך תושבי בני ברק שרובם חרדים. לקהילה זו דמוגרפיה ואורחות חיים ייחודיות ושונות מכלל אוכלוסיית ישראל, לכן ממצאי המחקר שמבוססים על קהילה זו עשויים להיות שונים מתמונת המצב באוכלוסייה הכללית.
- מדובר בעבודה המבוססת על נתונים תצפיתיים ולא על ניסוי מתוכנן ומוקפד. עובדה זו יכולה ליצור הטיות בנתונים ולכן לגרום להערכות להיות מוטות.
- מאפיינים חשובים כגון האם אדם הוא סימפטומטי, תאריכי הדבקה או הופעת הסימפטומים מתוך החקירות האפידמיולוגיות, מבוססים על דיווח עצמי ובשל כך עשויים להיות מוטות.
- הקריטריונים לביצוע בדיקה בקרב כל בני הבית אינם ידועים.
- הערכות המספריות לפרמטרים מבוססות על התאמת מודל מתמטי לנתונים אמפיריים ולכן אי הודאויות הכרוכות בהנחות המודל יכולות להשפיע על התוצאות.
- מדובר בעבודה ראשונית ויש צורך בהמשך מעקב וביצוע ניתוחים נוספים בכלל ולאחר פתיחת מוסדות החינוך בפרט.

### מסקנות והשלכות:

במחקר זה נעשה שימוש בשני מודלים. מודל ההדבקות במשפחה מצביע על כך שסיכוי ההידבקות (susceptibility) של ילדים נמוך מזה של מבוגרים - 20-50% ביחס למבוגרים (כאשר ילדים צעירים נדבקים פחות). אותו מודל מצביע על כך שסיכוי הילדים להדביק (infectivity) הוא 25-75% ביחס למבוגרים (עם תלות נמוכה בגיל הילד). המודל השני (הבינומי) מצביע גם הוא על כך שסיכוי ההידבקות (susceptibility) של ילדים נמוכים מאלו של המבוגרים, אך אינם זניחים.

בכל הנוגע למערכת החינוך שאלה מרכזית היא באיזו מידה הדבקה בתוך מוסדות החינוך תשפיע על קצב התפשטות המגפה הכולל בארץ. על מנת לענות על שאלה זו יש צורך להתייחס למאפייני ההדבקה, למאפיינים חברתיים של מוסדות החינוך (ריבוי מפגשים בין ילדים), ולמידת היישום של כללי הריחוק החברתי ודרישות שמירת ההיגיינה במוסדות אלו. בהתחשב בכל אלו פוטנציאל ההשפעה של פתיחת מערכת החינוך על התפשטות המגיפה עשוי להיות משמעותי, אך לא ניתן להעריכו בצורה וודאית בשלב זה.

### המלצות:

היות ופוטנציאל ההשפעה של פתיחת מערכת החינוך על התפשטות המגיפה עשוי להיות משמעותי, אנו ממליצים למקבלי ההחלטות לנהוג במשנה זהירות, ובשלב זה לבצע את פתיחת מערכת החינוך והחזרה ללימודים, באופן מדורג ובפריסה על פני זמן תוך כדי ניטור ובקרה כמומלץ בהמשך.

מאחר והמסקנות לגבי היכולת של ילדים להדביק אחרים הן ראשוניות בלבד, יש להמשיך לעקוב אחר הנושא באופן שיטתי. לפיכך, פתיחת מערכת החינוך צריכה להיות מלווה במערך ניטור ובקרה. בפרט, מומלץ להקים מערך שיאפשר ניטור ומעקב אחר מצב התחלואה במוסדות החינוך מרגע פתיחתם, חלקם או כולם. הניטור יאפשר זיהוי מיידי של מוסדות חינוך בהם ימצא חולה חיובי מקרב התלמידים או צוות העובדים. לאחר זיהוי חולה במוסד חינוך כלשהו, מומלץ לבצע מעקב אחר מצב ההדבקות במוסד ובקרב משפחות החולים ולנקוט צעדי מניעה להתפשטות ההדבקה, עד כדי סגירת המוסד/ות או מערכת החינוך כולה. לשם כך יש לבנות תכנית שכוללת הגדרת קריטריונים לרמת ההתערבות הנדרשת. יתר על כן, מומלץ לבצע בדיקות סקר במוסדות החינוך באופן יזום על מנת לקבל תמונת מצב עדכנית ולמנוע הדבקה והתפשטות המגפה דרך מוסדות החינוך. להקמת



מערך הניטור והבקרה מטרה נוספת והיא להבין את מאפייני ההדבקה של ילדים/מבוגרים ולתכנן מדיניות משופרת מבוססת נתונים להתמודדות עם המגפה כדוגמת המערך שהוקם באוסטרליה (ראה נספח 5).



## Household infection models and fitting to data

### 1 Model I: assuming no secondary infection

We first make the assumption that, in each household, all the infected members were infected by the index case (the first case, who was infected from the outside), with no secondary infections within the household. We then obtain a simple model that can be used to estimate the important parameters. We assume that in a household with an adult index case, the probability that a child will be infected is  $p_{ca}$  and the probability that an adult will be infected is  $p_{aa}$ . Similarly in a household with a child index case, the probability that a child will be infected is  $p_{cc}$  and the probability that an adult will be infected is  $p_{ac}$ . In this simple model, in a household with  $N_a$  adults and  $N_c$  children and an adult index case, the number of children infected in a household will be binomially distributed with  $n = N_c$ ,  $p = p_{ca}$ , and the number of adults infected (excluding the index case) will be binomially distributed with  $n = N_a - 1$  and  $p = p_{aa}$ , and similarly for a household with a child index case.

The parameters are easily estimated as the ratio of the number of individuals infected to the number of individuals at risk.  $p_{ac}$  ( $p_{ca}$ ) is estimated as the ratio of the number of adults (children) infected in households with a child (adult) index case to the total number of adults (children) in such households. For cutoff age 10 (i.e. individuals of age 10 and higher are considered adults), we obtain (with Clopper-Pearson binomial confidence intervals):

$$p_{ac} = 13/29 = 0.448 (0.264, 0.643), \quad p_{ca} = 135/698 = 0.193 (0.165, 0.225),$$

and  $p_{cc}$  ( $p_{aa}$ ) is estimated as the ratio of the number of adults (children) infected in households with an adult (child) index case, excluding the index case, to the total number of adults (children) in such households. We obtain:

$$p_{aa} = 637/1520 = 0.419 (0.394, 0.444), \quad p_{cc} = 4/20 = 0.2 (0.057, 0.437)$$

Assuming the cutoff is at age 20 (i.e. individuals of age 20 and higher are considered adults) the estimates are:

$$p_{cc} = 99/273 = 0.363 (0.305, 0.423), \quad p_{ca} = 302/1149 = 0.263 (0.238, 0.289)$$

$$p_{ac} = 63/136 = 0.463 (0.377, 0.551), \quad p_{aa} = 321/703 = 0.457 (0.419, 0.494)$$





## 2 Model II: a household outbreak

We now remove the assumption that the index case is the only infector in the household, so that we need to consider a household outbreak.

Assume a household has  $N_a$  adults and  $N_c$  children,  $N = N_a + N_c$ . The dynamic model we use for describing an outbreak in a household is a stochastic age-of-infection model, similar to the chain-binomial models traditionally used for household outbreaks ([1], ch. 14, [2], ch. 10), but somewhat more realistic in taking into account the generation-time distribution. The model is described below.

$S_a(t), S_c(t)$  denote the number of adults and children who are susceptible on day  $t$ .

$i_a(t), i_c(t)$  denote the number of adults and children who become infected on day  $t$ .

$P_\tau$  ( $1 \leq \tau \leq T$ ) denotes the generation-time distribution - it is fixed, estimated based on previous data.

$\beta_{ij}$  ( $i, j = a, c$ ) are transmission coefficients from individuals of type  $j$  to individuals of type  $i$ . These parameters will be fitted.

The dynamics is described by

$$\begin{aligned} i_a(t) &\sim \text{Bin} \left( S_a(t), \sum_{\tau=1}^T P_\tau [\beta_{aa} i_a(t-\tau) + \beta_{ac} i_c(t-\tau)] \right), \\ i_c(t) &\sim \text{Bin} \left( S_c(t), \sum_{\tau=1}^T P_\tau [\beta_{ca} i_a(t-\tau) + \beta_{cc} i_c(t-\tau)] \right), \\ S_a(t+1) &= S_a(t) - i_a(t), \quad S_c(t+1) = S_c(t) - i_c(t). \end{aligned}$$

The initial conditions are:

$$i_a(t) = i_c(t) = 0, \quad t < 0,$$

and, if the index case is an adult,

$$\begin{aligned} i_a(0) &= 1, \quad i_c(0) = 0 \\ S_a(0) &= N_a - 1, \quad S_c(0) = N_c, \end{aligned}$$

while if the index case is a child:

$$\begin{aligned} i_a(0) &= 0, \quad i_c(0) = 1 \\ S_a(0) &= N_a, \quad S_c(0) = N_c - 1. \end{aligned}$$

The household outbreak ends at time  $t^*$  if

$$i_a(t) = i_c(t) = 0, \quad t^* - T \leq t \leq t^* - 1.$$

The total number of infected among adults and children in the household is

$$I_a = N_a - S_a(t^*), \quad I_c = N_c - S_c(t^*).$$



### 3 Computing likelihood and fitting the model

Assume that we have information on a set of households. For each household we have the data:

$$(N_a, N_c, ind, I_a, I_c)$$

$N_a, N_c$  are the number of adults and children in the household,  $ind = a, c$  indicates whether the index case is an adult or a child.  $I_a, I_c$  are the number of adults and of children who were infected.

A household of type  $(N_a, N_c, ind)$  consists of  $N_a$  adults and  $N_c$  children, with an index case of type  $ind$  (adult or child). We will now look at the data for households of type  $(N_a, N_c, ind)$  and compute a log-likelihood  $LL_{N_a, N_c, ind}(\beta)$ , where  $\beta = (\beta_{aa}, \beta_{ac}, \beta_{ca}, \beta_{cc})$ . The log-likelihood corresponding to the entire data set will be the sum of the log-likelihoods corresponding to each household type:

$$LL(\beta) = \sum_{N_a, N_c, ind} LL_{N_a, N_c, ind}(\beta).$$

The log-likelihood  $LL_{N_a, N_c, ind}(\beta)$  is computed by a simulation approach, since we cannot compute it analytically (our approach is thus what is known as ‘simulated maximum likelihood’). We first run  $S$  (e.g.  $S = 1000$ ) realizations of the simulation of the household outbreak model for a household of type  $(N_a, N_c, ind)$ , with the given values for  $\beta$ . We generate a table of  $P(I_a, I_c)$  - the fraction of the simulations in which  $I_a$  adults and  $I_c$  children were infected. This is our estimate of the probability of the outcome  $(I_a, I_c)$  in a household of the above type.

Assuming the households of type  $(N_a, N_c, ind)$  in our data are indexed by  $j$ , with the outcome for household  $j$  given by  $(I_a(j), I_c(j))$ , the log-likelihood of this outcome is:

$$LL_{N_a, N_c, ind}(\beta) = \log \left( \prod_j P(I_a(j), I_c(j)) \right) = \sum_j \log P(I_a(j), I_c(j))$$

If we have  $M$  household types, computing  $LL(\beta)$  for each value of  $\beta$  will require  $M \cdot S$  simulations of the household model.

To estimate the parameters  $\beta$  we maximize  $LL(\beta)$  over the parameters  $\beta$ .

### 4 Special structure of the $\beta$ matrix - susceptibility and infectivity of children relative to adults

Under the assumption that adults and children differ in susceptibility to being infected and in degree of infectivity, we can write

$$\begin{aligned} \beta_{ac} &= \beta_{aa} \cdot \delta \\ \beta_{ca} &= \beta_{aa} \cdot \gamma, \\ \beta_{cc} &= \beta_{aa} \cdot \gamma \cdot \delta \end{aligned}$$



where  $\gamma$  is the susceptibility of children relative to that of adults, and  $\delta$  is the infectivity of children relative to that of adults.

While it is possible to estimate the parameters  $\beta_{aa}, \gamma, \delta$  by fitting the entire data set, we use an alternative approach which is computationally less demanding. We first estimate  $\beta_{aa}$  by fitting the model only to households consisting only of adults. Then, fixing  $\beta_{aa}$ , we estimate  $\beta, \gamma$  by fitting the model to the set of households which have children. In both stages we use the maximum-likelihood method described above.

## 5 Model fitting results

### 5.1 The data

Our data set contains information about 562 families. Data used includes age composition of families, ages of individuals who were tested positive and negative, and age of the index case.

To fit the model we must decide on a cutoff age between ‘children’ and ‘adults’. We fit the model several times, with different cutoff ages.

### 5.2 Parameter estimates

As described above, the parameter  $\beta_{aa}$  is estimated based on families containing only adults. Since the definition of ‘adult’ varies according to the cutoff, this also leads to variations in the value of  $\beta_{aa}$ . Figure 1 shows the graph of the negative of the log-likelihood as a function of  $\beta_{aa}$ . The maximum likelihood estimate is given by the minimizer of this function.

The parameters of greatest interest are the relative susceptibility  $\gamma$  and the relative infectivity  $\delta$ . Figure 2 shows the level sets of the negative log-likelihood as a function of the parameters  $\gamma, \delta$ . The maximum likelihood estimates for  $\gamma, \delta$  is given by the minimizer of this function.

## 6 Testing goodness of fit and comparing models

After fitting both models to the available data, we evaluate the goodness of fit. We do this by comparing aggregate outcomes of the fitted models to those in the data as follows: for each household size, we computed the distribution of the number of infected cases in a household according to the models, and compared them to the corresponding distributions obtained from the households in our data. These comparisons displayed in figure 3 (left). Visual inspection of fig 3 left, where the fits of the two models are presented, suggests that when there are not enough households of a certain size in the data ( $< 40$ ), the histograms are noisy and therefore it is difficult to compare the models fits. When the number of families (per family type) is larger than 40, it seems that Model II better fits the data. In addition, the right part of figure 3 displays the Bhattacharyya distance, measuring the similarity of the distributions obtained from the model

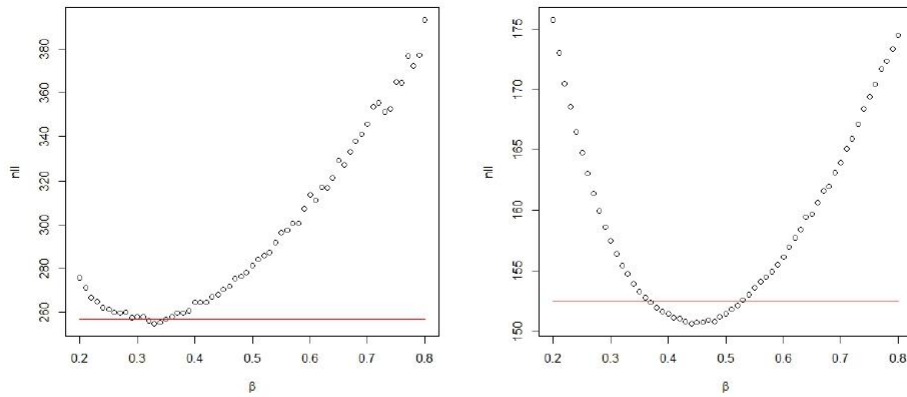


Figure 1: Graphs of the negative log-likelihood function for estimating  $\beta_{aa}$  based on households without children, where the cutoff age is 10 (left), 20 (right). Red line is 3 units above minimal value in the graph, used to obtain confidence intervals.

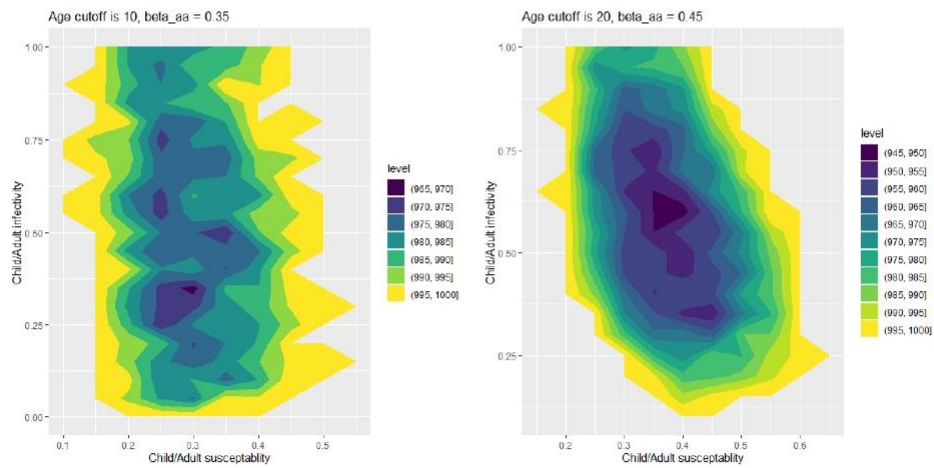


Figure 2: Level sets of log-likelihood function for estimating the relative infectivity  $\delta$  and the relative susceptibility  $\gamma$  of children, assuming cutoff age is 10 (left), 20 (right)

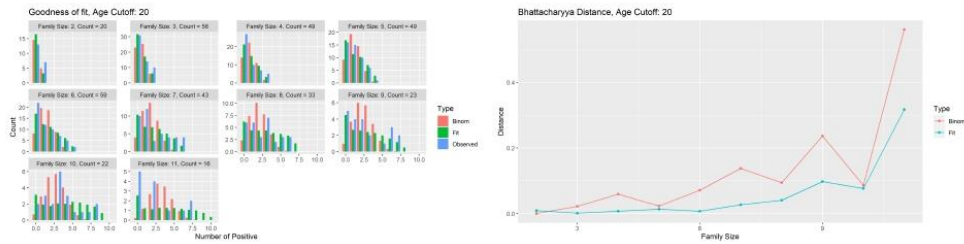


Figure 3: Comparison of distribution of the number of infected cases in households of different sizes in the data to that obtained from the two fitted models, for cutoff age 20 between children and adults.

to those obtained from the data. Here we see somewhat better fit of Model II - however, this requires further research and statistical validation.

## References

- [1] Bailey, N. T. (1975). The mathematical theory of infectious diseases and its applications. Charles Griffin & Company.
- [2] Becker, N. G. (2015). Modeling to inform infectious disease control (Vol. 74). CRC Press.



נספח 4: תיאור הנתונים

התפלגות גדלי המשפחות במדגם (עם חולה חיובי אחד לפחות במשפחה):

גודל בית אב	מספר משפחות	אחוז משפחות
1	157	21.0
2	156	20.9
3	84	11.3
4	64	8.6
5	54	7.2
6	62	8.3
7	49	6.6
8	33	4.4
9	28	3.8
10	21	2.8
11	19	2.5
12	7	0.9
13	8	1.1
14	2	0.3
15	2	0.3
סה"כ	<b>746</b>	<b>100</b>

התפלגות גדלי המשפחות שנבדקו במלואן ואין בהן חולים:

גודל בית אב	מספר משפחות	אחוז משפחות
1	1151	77.5
2	184	12.4
3	69	4.6
4	38	2.6
5	17	1.1
6	7	0.5
7	4	0.3
8	5	0.3
9	4	0.3
10	1	0.1
11	3	0.2
12	2	0.1
סה"כ	<b>1485</b>	<b>100</b>



התפלגות גדלי המשפחות שנבדקו אך לא במלואן (5,098):

גודל בית אב	כמות משפחות	כמות נדגמים (אחוז ממשפחות)	כמות חיוביים (אחוז מנדגמים)
2	665	665 (50%)	45 (6.8%)
3	691	817 (39.4%)	84 (10.3%)
4	701	906 (32.3%)	87 (9.6%)
5	677	967 (28.6%)	108 (11.2%)
6	581	873 (25%)	114 (13.1%)
7	483	723 (21.4%)	93 (12.7%)
8	395	715 (22.6%)	123 (17.2%)
9	327	631 (21.4%)	134 (21.6%)
10	212	413 (19.5%)	75 (18.2%)
11	124	253 (18.5%)	46 (18.2%)
12	95	194 (17%)	26 (13.4%)
13	61	163 (20.6%)	29 (17.8%)
14	39	99 (18.1%)	21 (21.2%)
15	20	40 (13.3%)	11 (27.5%)
16+	27	110 (21.4%)	12 (10.9%)



**סיבות לביצוע בדיקות קורונה:**

- מתוך 3,143 נבדקים במשפחות שנבדקו במלואן עם חיובי אחד לפחות, ל-1066 (34%) לא צויינה סיבת בדיקה, 990 מתוכם היו שליליים (92.8%).
- מתוך 2176 נבדקים במשפחות שנבדקו במלואן ולא נמצא בהן חולה, ל-1558 לא צויינה סיבת בדיקה.
- מתוך 7534 נבדקים במשפחות שלא נבדקו במלואן, ל-4499 אין סיבת בדיקה (מהם 4440 שליליים).
- מתוך המשפחות שנבדקו במלואן והיה בהן חולה אחד לפחות, להלן התפלגות הסיבות לבדיקה:

חול"ל	מגע עם חולה מאומת	סימפטומים	כמות במדגם	כמות בישראל
X	X	X	1066 (33.9%)	210813 (71.3%)
V	X	X	118 (3.8%)	5750 (1.9%)
V	V	X	0 (0%)	458 (0.15%)
V	X	V	391 (12.4%)	29407 (9.9%)
X	V	X	202 (6.4%)	2663 (0.9%)
X	V	V	690 (22%)	9649 (3.3%)
X	X	V	652 (20.7%)	35004 (11.8%)
V	V	V	24 (0.8%)	1925 (0.7%)

ובהסתכלות רק על חיוביים (1550 אנשים):

חול"ל	מגע עם חולה מאומת	סימפטומים	כמות	כמות בישראל
X	X	X	76 (4.9%)	1474 (9.7%)
V	X	X	8 (0.5%)	224 (1.6%)
V	V	X	0 (0%)	75 (0.5%)





1388 (9.2%)	53 (3.4%)	V	X	V
1559 (10.3%)	202 (13%)	X	V	X
5873 (38.8%)	689 (44.5%)	V	V	X
4148 (27.4%)	499 (32.2%)	V	X	X
386 (2.5%)	23 (1.5%)	V	V	V

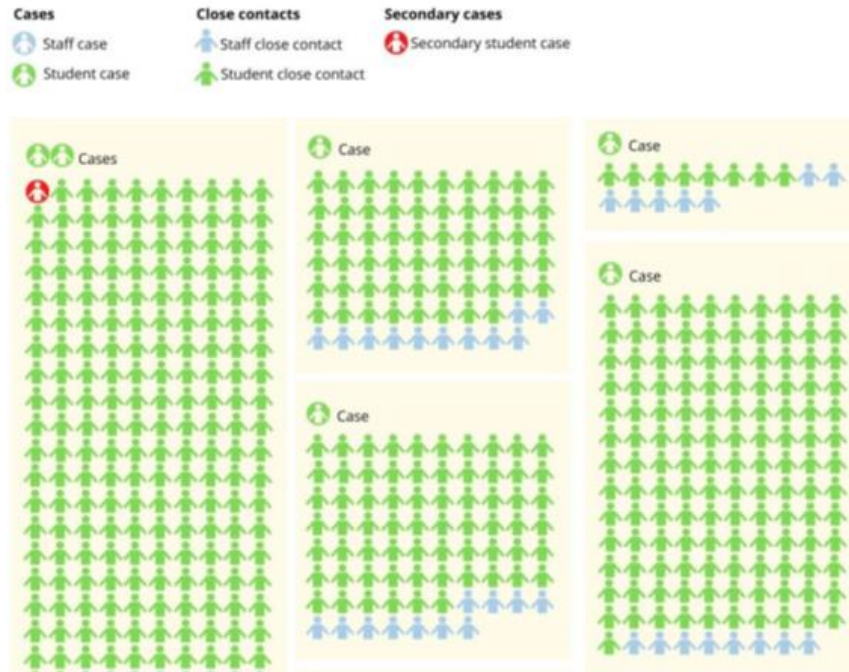
מתוך המשפחות שנבדקו במלואן והיה בהן לפחות חולה אחד, להלן אחוזי החולים לפי קבוצת גיל:

אסימפטומטיים (אחוז מחולים)	חולים (אחוז מסה"כ)	סה"כ	קבוצת גיל
50 (31.8%)	157 (19.6%)	802	0-9
97 (29.3%)	331 (42.2%)	785	10-19
139 (12.9%)	1077 (69.2%)	1556	20+



**נספח 5: מערך ניטור ובקרה באוסטרליה**

במחקר מדעי מאוסטרליה שפורסם השבוע, ניתן לראות כי מערך הבדיקה שלהם לא מסתפק בבידוד המגעים אלא ממשיך לעקוב אחריהם. הגרף הבא (איור 5) מראה את כמות המגעים של התלמיד החולה בכל כיתה, את מספר המגעים שהיו לו ואת מספר המודבקים.



לקוח מ:

COVID-19 in schools – the experience in NSW, Prepared by the National Centre for Immunisation Research and Surveillance 2020 (NCIRS) 26 April



## נספח 6: סקירת ספרות

Bi, Q., Wu, Y., Mei, S., Ye, C., Zou, X., Zhang, Z., ... & Gao, W. (2020). Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. MedRxiv.

נבדקו 391 מקרים בשנזן (סין) מאמצע ינואר עד אמצע פברואר, ו-1286 אנשים שהיו במגע קרוב עם מקרים אלו. מתוכם 20 ילדים בגיל 0-9. נמצא שאחוז ההדבקה של ילדים שהיו במגע עם חולה דומה לאחוז בכלל האוכלוסיה (כ 7%). הסיכוי להדבקה במגע עם חולה בתוך המשפחה היה 15% (לא צוין אחוז לפי גיל) אין התייחסות לסיכוי של ילד להדביק אחרים חלק משמעותי מהילדים היו א-סימפטומטיים או בעלי סימפטומים קלים.

Dong, Y., Mo, X., Hu, Y., Qi, X., Jiang, F., Jiang, Z., & Tong, S. (2020). [Epidemiology of COVID-19 among children in China](#). *Pediatrics*.

נבדקו 2135 ילדים חולים ברחבי סין במהלך ינואר-פברואר 2020. 34.1% אובחנו עם קורונה בבדיקת מעבדה והיתר הוגדרו חשודים בקורונה. מבין המקרים שאובחנו בבדיקת מעבדה, רב המקרים היו קלים 12.9% אסימפטומטיים, 43.1% קלים, 40.9% בינוני, 3% קשה וחמור 57% מהילדים שאובחנו היו בנים לא ניתן להסיק מהמחקר על סכויי ההדבקה של ילדים לעומת מבוגרים ועל הדבקה במגע עם ילד חולה

COVID, CDC, et al. "[Coronavirus Disease 2019 in Children—United States, February 12–April 2, 2020](#)." *Morbidity and Mortality Weekly Report* 69.14 (2020): 422.

נבדקו מקרים שדווחו למרכז לבקרת מחלות בארה"ב עד תחילת אפריל 2020, מתוכם 2531 ילדים. ברוב המקרים שדווחו חסר מידע מפורט על גיל, מצב רפואי קודם וסימפטומים. מהמקרים בהם דווח גיל וסימפטומים עולה כי 73% מהילדים היו א-סימפטומטיים. יתכן שתוצאה זו נובעת מהדיווח הלקוי.

Kelvin, A. A., & Halperin, S. (2020). COVID-19 in children: [the link in the transmission chain](#). *The Lancet Infectious Diseases*.

המחקר מנתח 36 מקרים של ילדים שאושפזו לאחר אבחון בקורונה. מתוכם 28% היו אסימפטומטיים (ילדים אלו נבדקו בגלל מגע קרוב עם חולה). הכותבים מסיקים שילדים יהיו לרב אסימפטומטיים או בעלי תסמינים קלים ולכן ישמשו חוליה משמעותית בהעברת המחלה

[Dutch National Institute for Public Health and the Environment](#)

המכון הלאומי לרפואה ציבורית בהולנד פירסם ממצאים של מחקרים בנושא הדבקה ילדים:



א. מחקר שנעשה בתוך משפחות – המחקר מצביע על כך שהילדים אינם המדביקים הראשונים בתוך המשפחה אך הם עשויים להידבק בתוך הבית.

ב. מחקר במרפאות בדק אנשים שהגיעו עם עם סימפטומים נשימתיים – באוכלוסיה הבוגרת נמצאו 6.5% - 30% חולי קורונה בשבועות שונים. לא נמצאו ילדים חולים בבדיקות אלו.

ג. ניתוח של מגעים עם חולים - ילדים שנדבקו לא הדביקו אחרים.

Lee, Ping-Ing, et al. "[Are children less susceptible to COVID-19?](#)" *Journal of Microbiology, Immunology, and Infection* (2020). מתוך 44672 מקרים שדווחו בסין עד תחילת פברואר 2020, רק 0.9% מהמקרים היו ילדים בגיל 0-9. הכותבים מסיקים שסכויי ההדבקה של ילדים נמוכים יותר ובוחנים סיבות אפשריות. המאמר לא שוקל את האפשרות לכך שמספר הילדים שאובחנו הוא נמוך משום שהם חולים קלים או א-סימפטומטיים.

Ludvigsson, J. F. (2020). [Systematic review of COVID-19 in children show milder cases and a better prognosis than adults](#). *Acta Paediatrica*.

המאמר מסכם 45 מחקרים שנעשו על התפרצות הקורונה ודיווחו על ילדים חולים. מרבית המידע הגיע מסין ומארה"ב. נמצא שילדים היוו 1%-5% מהמאובחנים במחקרים אלו:

בסין – 2% מהמאובחנים היו בגיל 0-19, 1% היו בגיל 0-9 באיטליה – 1.2% מהמאובחנים הם ילדים

בארה"ב – 5% מהמאובחנים הם ילדים, 1% מהמאושפזים הם ילדים במחקר העיקרי בסין 9% מהילדים היו א-סימפטומטיים, קלים או בינוניים המאמר מסיק שהילדים נוטים לחלות פחות או שהם א-סימפטומטיים/בעלי סימפטומים קלים ולכן לא מאובחנים

Cai, J., Xu, J., Lin, D., Xu, L., Qu, Z., Zhang, Y., ... & Xia, A. (2020). [A Case Series of children with 2019 novel coronavirus infection: clinical and epidemiological features](#). *Clinical Infectious Diseases*. המאמר מנתח בפרוט 10 מקרים של ילדים שנדבקו בסין. ב- 8 מהמקרים הילדים נדבקו ממבוגר בתוך המשפחה. ב 6 מקרים, נדבקו ילדים נוספים במשפחה, והמאמר מסיק שמקור ההדבקה הוא הילד החולה.

Gudbjartsson, D. F., Helgason, A., Jonsson, H., Magnusson, O. T., Melsted, P., Norddahl, G. L., ... & Eiriksdottir, B. (2020). [Spread of SARS-CoV-2 in the Icelandic Population](#). *New England Journal of Medicine*.

מחקר שבוצע באיסלנד בדק באופן ממוקד אוכלוסיה חשודה (בעלי סימפטומים שהיו במגע עם חולה) ובנוסף ביצע שני סקרים באוכלוסיה הכללית: אחד לאנשים שהביעו ענין להיבדק והשני לאוכלוסיה אקראית. בסקרי האוכלוסיה נמצא כי אחוז החולים המאובחנים בכלל האוכלוסיה הוא 0.6%-0.8%. בקרב ילדים עד גיל 10 לא נמצאו חולים בסקרי האוכלוסיה. בילדים מעל גיל 10 נמצא כי אחוז המאובחנים הוא 0.8%, כמו באוכלוסיה הכללית.

בבדיקה הממוקדת הסכוי של ילדים להימצא חולים היה כמחצית מהסכוי של מבוגרים (6.7% לעומת 13.7%)

[Australian National Center for immunization Research and surveillance](#)

מכון מחקר באוסטרליה תיעד 18 ארועים שבהם 9 תלמידים ו 9 אנשי סגל ובתי ספר באוסטרליה אובחנו כחולים. נמצא כי 735 תלמידים ו 128 חולים היו במגע קרוב עם לפחות חלק מהחולים. מתוכם – אף איש סגל לא נדבק. מבין התלמידים שהיו במגע עם החולים נדבק תלמיד אחד בבית ספר יסודי ואחד בבית ספר תיכון.



Davies, N. G., Klepac, P., Liu, Y., Prem, K., Jit, M., Eggo, R. M., & CMMID COVID-19 working group. (2020). [Age-dependent effects in the transmission and control of COVID-19 epidemics](#). *medRxiv*. החוקרים בנו מודל להתפשטות המגיפה והתאימו אותו לתוצאות בוואן. מהתאמת הפרמטרים הם מסיקים שלילדים מתחת לגיל 10 יש סכוי של 20% להראות סימפטומים להערכתם הם נדבקים באופן דומה למבוגרים. המחקר לא מבוסס על ניתוח של ממצאים קליניים.

Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W., Zhao, S., ... & Ajelli, M. (2020). [Age profile of susceptibility, mixing, and social distancing shape the dynamics of the novel coronavirus disease 2019 outbreak in China](#). *medRxiv*. החוקרים ניתחו 1245 מגעים שדווחו ע"י 636 חולים בוואן ו 1296 מגעים שדווחו ע"י 557 חולים בשנחאי. החוקרים מסיקים שהסיכוי של ילדים להידבק נמוך מזה של מבוגרים בכ- 60%. הסכוי לסימפטומים אצל ילדים, בהינתן הדבקה, נמוך בכ 30% בהשוואה למבוגרים.

השוואה בין המחקרים

מחקר	% א-סימפטומטיים/ סימפטומטיים בקרב ילדים	סכוי הדבקה אצל ילדים	הערות
Zhang et al.	סכוי לסימפטומים נמוך ב 30% בהשוואה למבוגר	נמוך ב 60% בהשוואה למבוגר	מחקר קליני
Davies et al.	80% אסימפטומטיים	דומה למבוגר	לא קליני
Gudbjartsson et al.	-	מחצית ממבוגר	קליני
Kelvin et al.	28% אסימפטומטיים	-	קליני
CDC	73% סימפטומטיים		מבוסס על דיווח חלקי
Dong et al.	12.9%		קליני
Bl et al.		דומה למבוגר	קליני