

Dokumentace k epidemiologickému modelu ÚZIS ČR pro krátkodobé predikce

Autoři:	Ondřej Májek, Ondřej Ngo, Jiří Jarkovský, Monika Ambrožová, Barbora Budíková, Jan Kouřil, Ladislav Dušek, Tomáš Pavlík
Verze:	1.0
Datum:	8. 6. 2020

1 Obsah

1	Obsah.....	2
2	Úvod	3
2.1	Využití epidemiologického modelu pro krátkodobé predikce	3
2.2	Podstata a vědecká relevance tohoto typu modelů	3
2.3	Metodické citace	3
3	Vstupní a výstupní údaje	4
3.1	Vysvětlení základních pojmů	4
3.2	Vstupy do použitého modelu	5
3.3	Výstupy modelu.....	6
4	Struktura modelu	7
5	Hodnoty parametrů z aktualizace dne 20.5.2020	8
6	Popis implementace výpočetního postupu	8
6.1	Rovnice přechodu.....	8
6.2	Výstupy modelu.....	8
7	Rešerše pro informování struktury a parametrů modelu	12
7.1	Reprodukční číslo	12
7.2	Inkubační doba	14
7.3	Sériový interval.....	16
7.4	Podíl asymptomatických/subklinických případů	18

2 Úvod

2.1 Využití epidemiologického modelu pro krátkodobé predikce

- Krátkodobá předpověď počtu nově potvrzených případů onemocnění
- Podklad pro předpovědi dalších klíčových charakteristik: počet hospitalizovaných, počet osob v těžkém stavu nebo s vysoce intenzivní péčí, počet úmrtí
- Vyhodnocení dynamiky epidemie, stanovení reprodukčního čísla, hodnocení souvislostí se zavedenými opatřeními

2.2 Podstata a vědecká relevance tohoto typu modelů

Jedná se o rozšíření epidemiologických SIR modelů, které jsou standardem v modelování vývoje akutních infekčních onemocnění¹.

Model reprezentuje dynamiku infekčního onemocnění v populaci. Jedinci v populaci jsou v každém okamžiku uvažováni v jednom ze stavů (kompartmentů), jedná se tedy v principu o tzv. stavový (kompartmentový) model S(E)IR:

- S náchylní jedinci
- I infikovaní jedinci
 - Prakticky se tento stav dělí v modelech COVID na
 - E latentní onemocnění, před infekčností
 - Vlastní I infekční jedinci
- R jedinci po konci infekčnosti

S modely je spjata parametrická a strukturní nejistota a nejistota ohledně budoucích opatření a jejich dopadu na dynamiku epidemie.

2.3 Metodické citace

Příklad recentních aplikací S(E)IR modelu v odborné literatuře

- WU, Joseph T.; LEUNG, Kathy; LEUNG, Gabriel M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet*, 2020, 395.10225: 689-697.
- KUCHARSKI, Adam J., et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The lancet infectious diseases*, 2020.
- JIA, Wangping, et al. Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China. *medRxiv*, 2020. (článek zatím neprošel recenzním řízením)

¹ LIU, Xinzhi; STECHLINSKI, Peter. *Infectious Disease Modeling*. Springer, 2017.

3 Vstupní a výstupní údaje

3.1 Vysvětlení základních pojmů

Inkubační doba

- představuje časový interval mezi infekcí a prvními příznaky onemocnění
- v modelovém hodnocení umožňuje modelovat dosud nepotvrzené případy nákazy a šíření nákazy v infekční fázi inkubační doby pacienta

Infekční doba

- představuje časový interval nakažlivosti infikovaného pacienta
- v modelovém hodnocení představuje nepostradatelnou informaci o časovém období možného šíření nákazy

Sériový interval

- představuje interval mezi nástupem příznaků u nakažlivého jedince a prvním nástupem příznaků u nakažené osoby
- v modelovém hodnocení představuje rychlost šíření nákazy mezi jedinci
- lze odvodit z inkubační a infekční doby

Podíl subklinických případů

- v našem modelu představuje infekci u pacienta, který nemá žádné či minimální příznaky onemocnění, ale může nakazit další osoby (definice se může mezi publikacemi lišit)

Efektivita testování

- odpovídá době od nástupu symptomů po potvrzení výsledku a zaznamenání do databáze ISIN
- v modelovém hodnocení ovlivňuje souvislost mezi průběhem onemocnění a reportováním do databáze ISIN

Importované případy

- představují potvrzené nákazy onemocnění u pacientů v ČR, kteří byli infikováni v zahraničí
- v modelovém hodnocení představují importované případy iniciální ohniska pro modelování šíření nákazy v ČR

3.2 Vstupy do použitého modelu

Parametr	Hodnota	Zdroj	Užití
Počty osob s prokázanou nákazou	Datová sada Celkový počet osob s prokázanou nákazou dle KHS/laboratoří	Informační systém infekčních nemocí	Cílová hodnota pro kalibraci modelu
Inkubační doba	6 dní	Literární přehled (viz kapitola 7.2)	zohlednění ve struktuře modelu
Infekční doba	3 dny	Předpoklad v předchozích publikacích, např. Kucharski et al., 2020 ²	zohlednění ve struktuře modelu
Sériový interval	5 dní	Literární přehled (viz kapitola 7.3)	soulad s údaji o inkubační a infekční době
Podíl subklinických případů	10 % (souběžně zkoumána i alternativa 30 %)	Předpoklad v předchozích publikacích, např. Hellewell et al., 2020 ³ (viz také literární přehled, kapitola 7.4)	Parametr modelu
Efektivita testování	Data IS IN	Informační systém infekčních nemocí	Parametr pro kalibraci
Počty importovaných případů	Datová sada Přehled osob s prokázanou nákazou dle hlášení KHS	Informační systém infekčních nemocí	Nastavení počáteční dynamiky epidemie
Věková struktura osob s prokázanou nákazou	Datová sada Přehled osob s prokázanou nákazou dle hlášení KHS	Informační systém infekčních nemocí	Odhad počtu seniorů

² KUCHARSKI, Adam J., et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The lancet infectious diseases, 2020.

³ HELLEWELL, Joel, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. The Lancet Global Health, 2020.

3.3 Výstupy modelu

Z modelu lze získat tzv. markovskou stopu, časově podmíněný záznam všech uvažovaných charakteristik. Jako klíčové výstupy modelu jsou uvažovány:

- Kumulativní počet potvrzených případů onemocnění COVID-19
- Počet nově potvrzených případů
- Aktuální počet dosud nepotvrzených infikovaných jedinců

Jako výstup modelu zároveň slouží kalibrované hodnoty parametrů modelu, zejména **základní/efektivní reprodukční číslo**, standardně označované jako R :

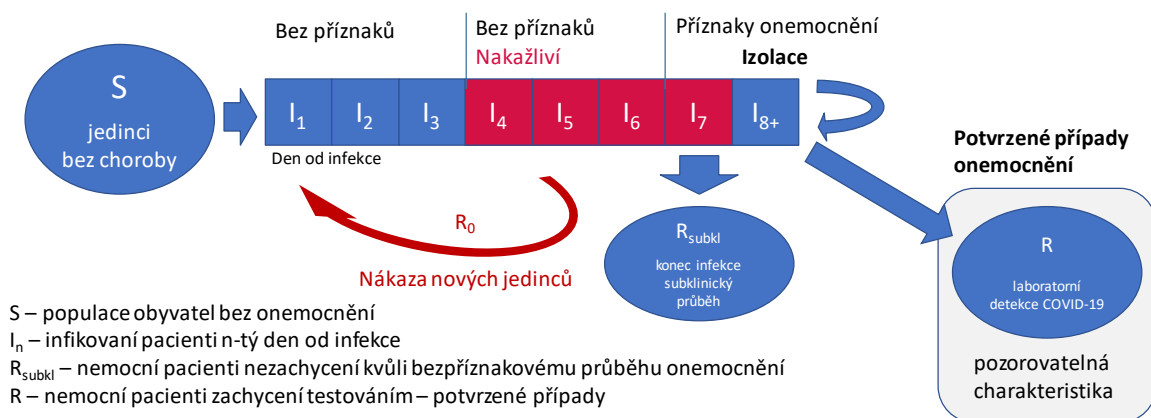
- představuje průměrný počet dalších osob, které přímo nakazí jeden infikovaný jedinec
- základní reprodukční číslo udává počáteční hodnotu v dané populaci před přijetím ochranných opatření, postupně by se mělo snižovat na tzv. efektivní reprodukční číslo, které odpovídá pomalejšímu šíření epidemie díky přijatým opatřením
- v modelovém hodnocení představuje klíčový parametr infekčnosti onemocnění a zároveň zohledňuje přijatá opatření v čase

Intervalový odhad efektivního reprodukčního čísla lze získat jako součást kalibračního procesu. Aplikací tohoto intervalového odhadu v pravděpodobnostní analýze senzitivity lze získat „interval senzitivity“ pro kumulativní počet případů k danému referenčnímu datu.

Výstupy modelu jsou zároveň konfrontovány s odbornou literaturou (viz literární přehled, kapitola 7.1).

4 Struktura a předpoklady modelu

Struktura modelu je vysvětlena na následujícím schématu



Model predikuje průchod pacientů průběhem onemocnění, s definovanou délkou inkubační doby. Noví pacienti přicházejí do modelu importem nebo nákazou, končí se subklinickým průběhem nebo jako potvrzený případ (**jediná přímo sledovaná charakteristika**). Klíčovým parametrem modelu je **tzv. reprodukční číslo**: průměrný počet osob, které nakazí 1 nakažená osoba.

Předpoklady modelu:

- do modelu vstupují importované případy (nakažení je datováno 6 dní před datem počátku symptomů dle ISIN, kalibrace provedena na konci března)
- setrvání ve stavech I_1 - I_7 trvá vždy právě 1 den, setrvání ve stavu I_{8+} je podmíněno efektivitou testování (pravděpodobnost přechodu, resp. průměrná doba setrvání, byla orientačně kalibrována na reálná data ISIN)
- počet nově nakažených (vstup do I_1) odpovídá reprodukčnímu číslu
- šestidenní inkubační doba (Hellewell et al., 2020), druhá polovina intervalu představuje infekční období (alternativní předpoklad z Kucharski et al., 2020, odpovídá sériovému intervalu 5 dní)
- druhý den po projevu příznaků izolace, předpoklad konce nakažlivosti pro populaci
- efektivita testování kalibrována na pozorované data, předpoklad postupného zlepšování efektivity (zkracování doby do diagnózy)
- model tedy předpokládá testování pouze osob s příznaky
- subklinický průběh: uvažováno 10 % (Hellewell et al., 2020), variantně i 30 %

5 Popis implementace výpočetního postupu

5.1 Implementace

Výpočet je implementován jako markovský model v programu MS Excel.

5.2 Výpočetní postup modelu

Stavy v čase $t+1$:

- $I_1(t+1) = \text{PŘÍJEZDY}(t+1) + (I_4(t) + I_5(t) + I_6(t) + I_7(t)) * \text{beta}(t+1)$
- $I_2(t+1) = I_1(t)$
- $I_3(t+1) = I_2(t)$
- $I_4(t+1) = I_3(t)$
- $I_5(t+1) = I_4(t)$
- $I_6(t+1) = I_5(t)$
- $I_7(t+1) = I_6(t)$
- $R_{\text{subkl}}(t+1) = I_7(t) * \text{subcl_prop} + R_{\text{subkl}}(t)$
- $R(t+1) = I_{8+}(t) * \text{eftest}(t+1) + R(t)$
- $I_{8+}(t+1) = I_7(t) + I_{8+}(t) - I_7(t) * \text{subcl_prop} - I_{8+}(t) * \text{eftest}(t+1)$

(počty pacientů v jednotlivých stavech jsou zaokrouhlovány na celá čísla)

Stav v čase $t = 0$:

- $I_1(0) = 0$
- $I_2(0) = 0$
- $I_3(0) = 0$
- $I_4(0) = 0$
- $I_5(0) = 0$
- $I_6(0) = 0$
- $I_7(0) = 0$
- $R_{\text{subkl}}(0) = 0$
- $R(0) = 0$
- $I_{8+}(0) = 0$

5.3 Výstupy modelu

- Kumulativní počet potvrzených případů
 - Je dán hodnotou $R(t)$
- Dosud nepotvrzení infikovaní jedinci
 - $I_1(t) + \dots + I_{8+}(t)$

5.4 Kalibrace modelu

Modul pro kalibraci modelu je naprogramován v jazyce Microsoft Visual Basic for Applications.

Kalibračním cílem je kumulativní počet potvrzených případů COVID-19 (v jednotlivých dnech od začátku epidemie). Metrikou shody modelu s pozorováním je součet čtverců odchylek $R(t)$ od kalibračního cíle. Parametr byl hledán metodou náhodného prohledávání prostoru daného přípustnými hodnotami kalibrovaných parametrů, viz tabulka v kapitole 6.

Akceptováno bylo 5 % iterací s nejlepší shodou.

Z tohoto souboru byly stanoveny odhady parametrů modelu (jejich průměr a směrodatná odchylka), včetně intervalového odhadu reprodukčního čísla od 1. 5. k použití do analýzy senzitivity.

6 Hodnoty parametrů z aktualizace dne 20.5.2020

Jméno parametru	Užitá hodnota	Rozsah pro kalibraci (20.5.2020)	
		min	max
beta1	0,66	Z minulých verzí modelu	
beta2	0,46		
beta3	0,32		
beta4	0,25		
beta5	0,18	0,175	0,185
beta6	0,27	0,150	0,350
eftest1	0,14	Z minulých verzí modelu	
eftest2	0,44	0,400	0,500
subcl_prop	0,10	Předpoklad	

Hodnoty **beta1** až **beta6** odpovídají čtvrtině příslušného reprodukčního čísla (počtu nakažených jedním infekčním jedincem každý den, za předpokladu infekčního období 4 dny) v příslušném časovém období.

Reprodukční číslo je odhadováno samostatně v segmentech se zlomovými body danými zaváděním vládních opatření, resp. počátky posledních kalendářních měsíců: 7. 3., 12. 3., 16. 3., 1. 4. a 1.5.

Hodnoty reprodukčního čísla:

- před 7.3.
 - celkové reprodukční číslo R_0 : 2,64
- od 7.3. (karanténa pro cestující z Itálie)
 - celkové reprodukční číslo R : 1,84
- od 12.3. (den po uzavření škol, nouzový stav)
 - celkové reprodukční číslo R : 1,28
- od 16.3. (omezení volného pohybu osob)
 - celkové reprodukční číslo R : 1,00
- od 1.4.
 - celkové reprodukční číslo R : 0,72 (95% interval spolehlivosti 0,715 – 0,727)
- od 1.5.
 - celkové reprodukční číslo R : 1,08 (95% interval spolehlivosti 0,87 – 1,29)

Efektivita testování byla kalibrována ve třech segmentech se zlomovými body 22. 3. a 1. 4. V prvním segmentu dosahuje pravděpodobnost přechodu $I_{8+} \rightarrow R$ 14 % (efest1), ve třetím 44 % (efest2), s lineárním růstem v prostředním segmentu.

Podíl subklinických případů (subcl_prop) je uvažován na úrovni 10 %.

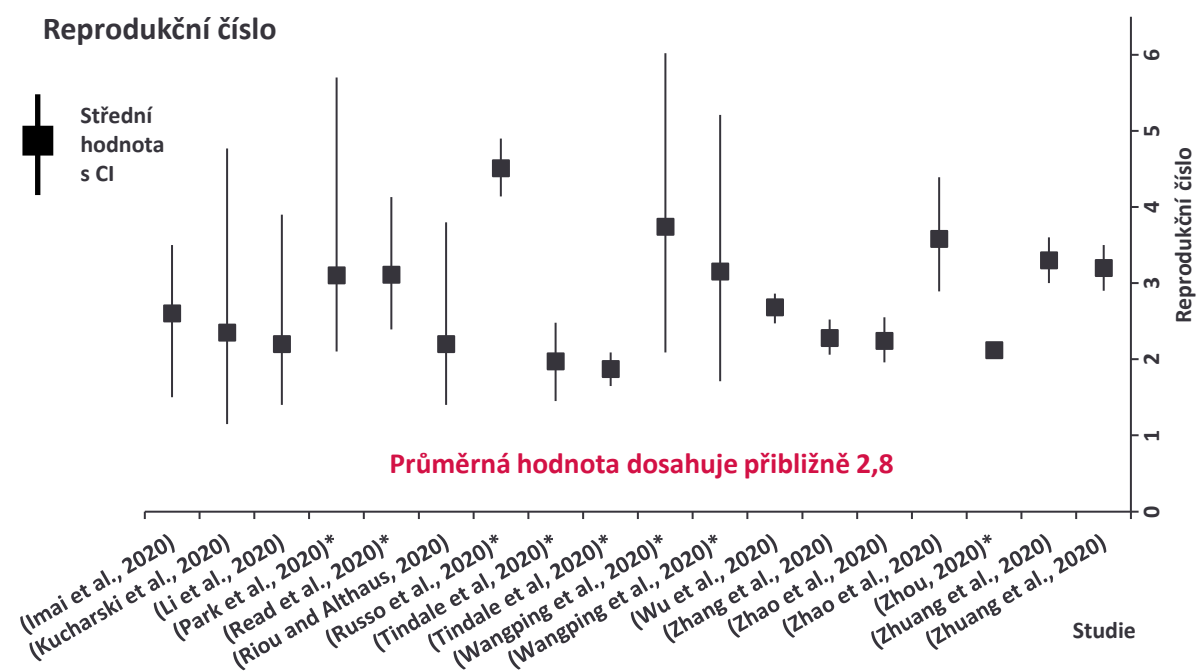
Součástí modelových parametrů je i časová řada PŘÍJEZDY(t), zahrnující **počty importovaných případů**. Hodnota základního reprodukčního čísla byla kalibrována s využitím počtů importovaných případů z IS IN dne 28.3.2020. Jako datum počátku infekce bylo zvoleno datum počátku příznaků (dle IS IN) zmenšené v souladu s předpoklady modelu o 6 dní. Další hodnoty byly extrapolovány s předpokladem příjezdů 3 osob denně.

Podíl seniorů (60+) je vypočítán z posledních sedmi dnů z dat IS IN. Retrospektivní počty nově diagnostikovaných seniorů jsou převzaty z dat IS IN.

7 Rešerše pro informování struktury a parametrů modelu

Rešerše aktualizována dne 24.5.2020. Rešerše představuje základní přehled publikované literatury pro informování dalšího vývoje modelů, není úplným systematickým přehledem.

7.1 Reprodukční číslo



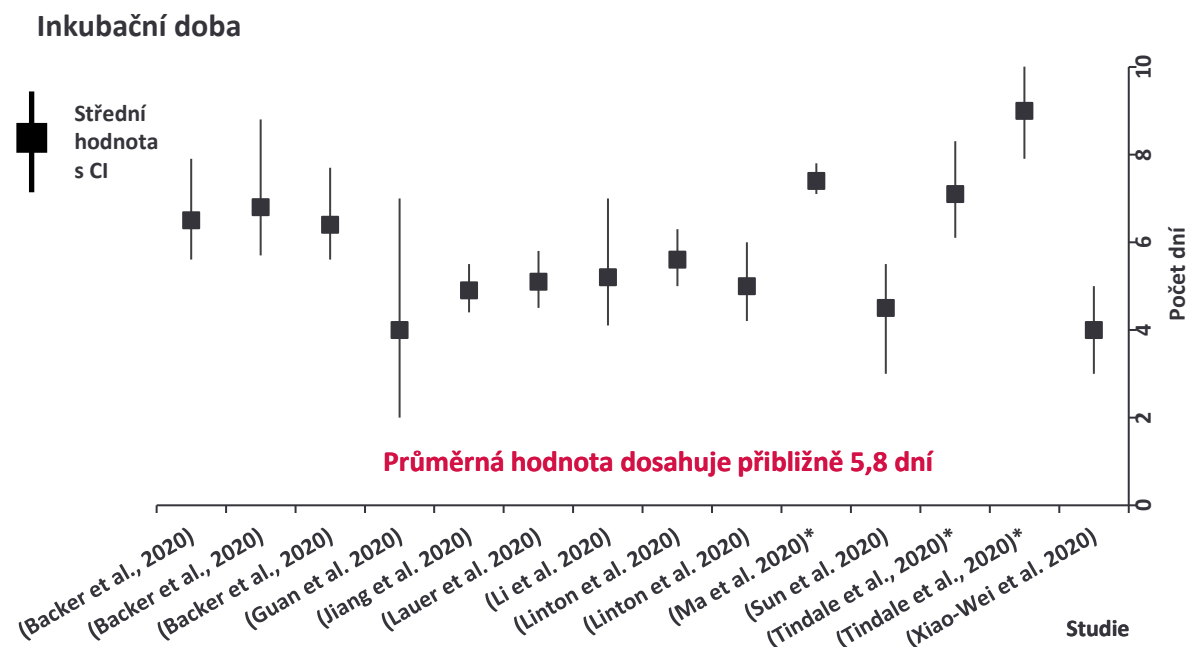
Zdroj	Sledovaná populace (N)	Sledované období	Výsledky Průměr (95% CI)
(Imai et al., 2020)	Wuhan, Čína (N cca 4 000)	do 18. 1. 2020	2,60 (1,50-3,50)
(Kucharski et al., 2020)	Wuhan, Čína	do 16. 1. 2020	2,35 (1,15-4,77)**
(Li et al., 2020)	Wuhan, Čína (N = 425)	do 22. 1. 2020	2,20 (1,40-3,90)
(Park et al., 2020)	Odhad pooled R0 ze 7 studií		3,10 (2,10-5,70)**
(Read et al., 2020)	Wuhan, Čína	1. 1. 2020 – 22. 1. 2020	3,11 (2,39-4,13)
(Riou and Althaus, 2020)	Čína	do 18. 1. 2020	2,20 (1,40-3,80)**

Zdroj	Sledovaná populace (N)	Sledované období	Výsledky Průměr (95% CI)
(Russo et al., 2020)	Lombardie, Itálie	21. 2. 2020 – 8. 3. 2020	4,51 (4,14-4,90)
(Tindale et al., 2020)	Singapur (N = 93)	19. 1. 2020 – 26. 2. 2020	1,97 (1,45-2,48)
(Tindale et al., 2020)	Tianjin, Čína (N = 135)	21. 1. 2020 – 27. 2. 2020	1,87 (1,65-2,09)
(Wangping et al., 2020)	Itálie	22. 1. 2020 – 16. 3. 2020	3,74 (2,09-6,02)
(Wangping et al., 2020)	Hunan, Čína	22. 1. 2020 – 16. 3. 2020	3,15 (1,71-5,21)
(Wu et al., 2020)	Wuhan, Čína	31. 12. 2019 – 28. 1. 2020	2,68 (2,47-2,86)
(Zhang et al., 2020)	Diamond Princess (N = 3 711)	do 16. 2. 2020	2,28 (2,06-2,52)**
(Zhao et al., 2020)	Čína	10. 1. 2020 – 24. 1. 2020	Od: 2,24 (1,96-2,55) Do: 3,58 (2,89-4,39)
(Zhou et al., 2020)	Wuhan, Čína	29. 1. 2020 – 2. 2. 2020	2,12 (2,04-2,18)
(Zhuang et al., 2020)	Itálie	5. 2. 2020 – 5. 3. 2020	3,30 (3,00-3,60)
(Zhuang et al., 2020)	Jižní Korea	20. 1. 2020 – 1. 3. 2020	3,20 (2,90-3,50)

* studie jsou v recenzním řízení

** výsledek zapsán ve formátu: medián (95% CI)

7.2 Inkubační doba



Zdroj	Sledovaná populace (N)	Sledované období	Výsledky Průměr dny (95% CI)
(Backer et al., 2020)	Wuhan, Čína (N = 88)	20. 1. – 28. 1. 2020	6,4 (5,6-7,7) 6,5 (5,6-7,9) 6,8 (5,7-8,8)
(Guan et al., 2020)	Čína (N = 291)	11. 12. 2019 – 29. 1. 2020	4,0 (2,0-7,0)**
(Jiang et al., 2020)	Čína (N = 50)	k 8. 2. 2020	4,9 (4,4-5,5)
(Lauer et al., 2020)	Čína (N = 181)	4. 1. – 24. 2. 2020	5,1 (4,5-5,8)**
(Q. Li et al., 2020a)	Wuhan, Čína (N = 10)	do 22. 1. 2020	5,2 (4,1-7,0)
(Linton et al., 2020)	Čína (N = 158; N = 52)	do 31. 1. 2020	5,6 (5,0-6,3)

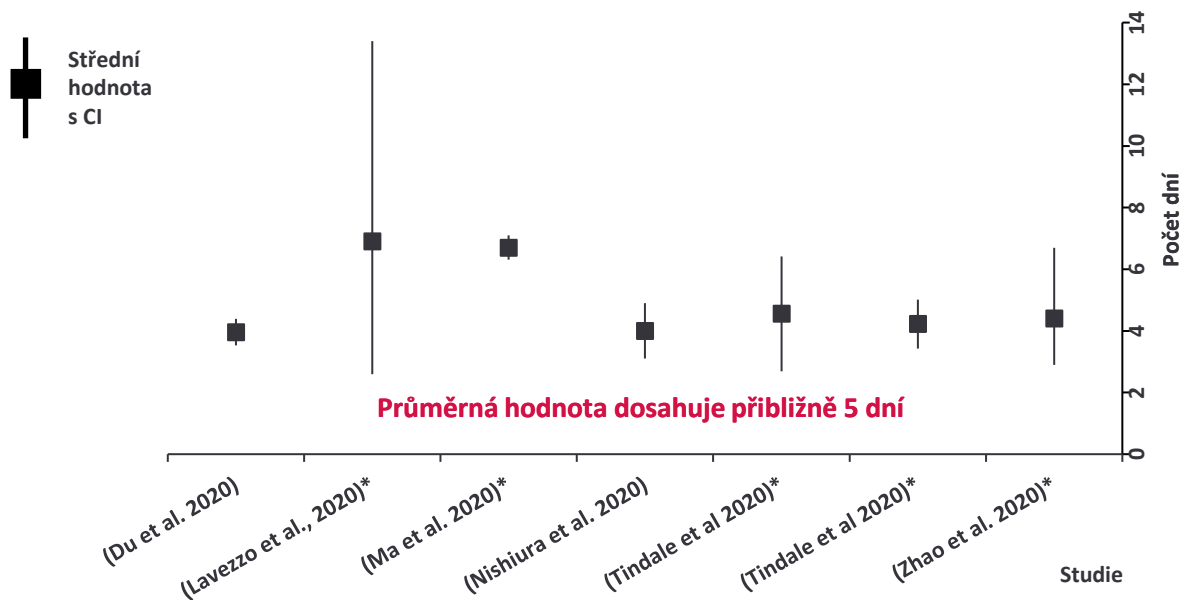
Zdroj	Sledovaná populace (N)	Sledované období	Výsledky Průměr dny (95% CI)
			5,0 (4,2-6,0)
(Ma et al., 2020)*	Čína, Japonsko, Singapur, Německo, Jižní Korea, Vietnam, Malajsie (N = 587)	konec února – začátek března	7,4 (7,1-7,8)
(Sun et al., 2020)	Čína (N = 507)	13. 1. – 31. 1. 2020	4,5 (3,0-5,5)**
(Tindale et al., 2020)*	Singapur (N = 93)	19. 1. 2020 – 26. 2. 2020	7,1 (6,1-8,3)
(Tindale et al., 2020)*	Tianjin, Čína (N = 135)	21. 1. 2020 – 27. 2. 2020	9,0 (7,9-10,2)
(Xu et al., 2020)	Zhejiang, Čína (N = 56)	10. 1. – 26. 1. 2020	4 (3-5)**

* studie jsou v recenzním řízení

** výsledek zapsán ve formátu: medián (95% CI)

7.3 Sériový interval

Sériový interval



Zdroj	Sledovaná populace (N)	Sledované období	Výsledky Průměr dny (95% CI)
(Du et al., 2020)	Čína (N = 468)	28. 1. – 8. 2. 2020	3,96 (3,53-4,39)
(Lavezzo et al., 2020)*	Vo, Itálie (2 812; 2 343)	23. 2. – 8. 3. 2020	6,9 (2,6-13,4)
(Ma et al., 2020)*	Čína, Japonsko, Singapur, Německo, Jižní Korea, Vietnam, Malajsie (N = 689)	konec února – začátek března	6,7 (6,31-7,10)
(Nishiura et al., 2020b)	Německo (N = 28)		4,0 (3,1-4,9)**
(Tindale et al., 2020)*	Singapur (N = 93)	19. 1. 2020 – 26. 2. 2020	4,56 (2,69-6,42)

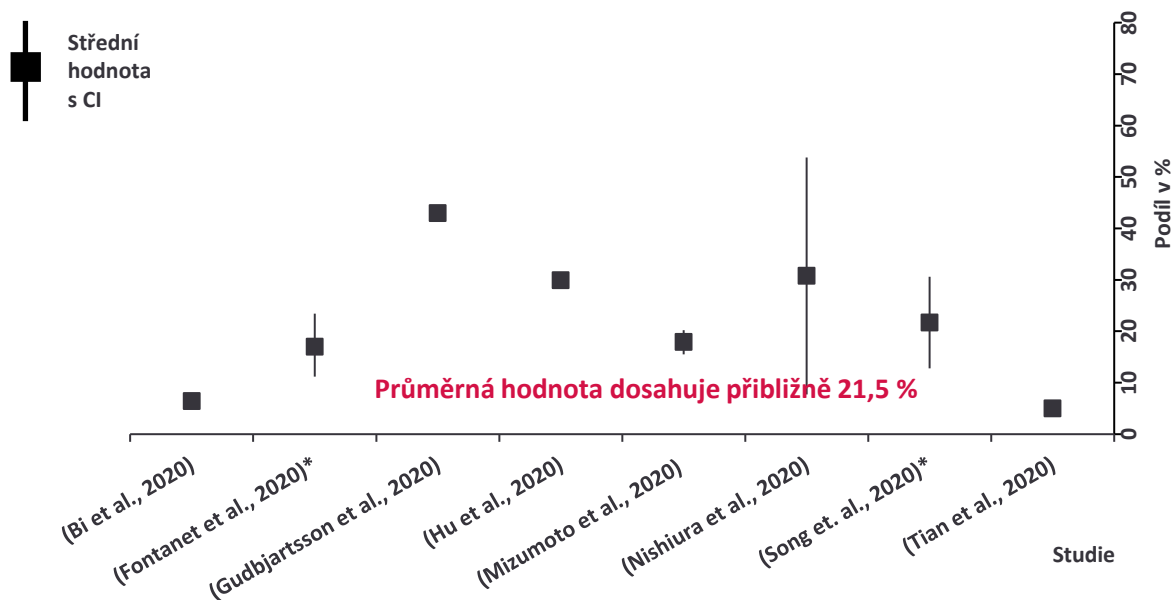
Zdroj	Sledovaná populace (N)	Sledované období	Výsledky Průměr dny (95% CI)
(Tindale et al., 2020)*	Tianjin, Čína (N = 135)	21. 1. 2020 – 27. 2. 2020	4,22 (3,43-5,01)
(Zhao et al., 2020a)*	Hong Kong (N = 21)	16. 1. – 15. 2. 2020	4,4 (2,9-6,7)

* studie jsou v recenzním řízení

** výsledek zapsán ve formátu: medián (95% CI)

7.4 Podíl asymptomatických/subklinických případů

Podíl asymptomatických/subklinických případů



Zdroj	Sledovaná populace (N)	Sledované období	Výsledky Podíl (95% CI)
(Bi et al., 2020)	Shenzen, Čína (N = 391)	14. 1. – 12. 2. 2020	6,4 %
(Fontanet et al., 2020)*	Střední škola v Oise, Francie (N = 661)	30. 3. – 4. 4. 2020	17,0 % (11,2–23,4)
(Gudbjartsson et al., 2020)	Island (N = 13 080)	31. 1. – 4. 4. 2020	43 %
(Hu et al., 2020)	Čína (N = 24)		29,9 %
(Mizumoto et al., 2020)	Posádka lodi Diamond Princess (N = 3 711)	do 20. 2. 2020	17,9 % (15,5–20,2)
(Nishiura et al., 2020a)	Evakuační let osob z Číny do Japonska (N = 13)	k 6. 2. 2020	30,8 % (7,7–53,8)

(Song et al., 2020)*	Ganzi (Tibet, Čína) (N = 83)	26. 1. – 6. 3. 2020	21,7 % (12,8-30,6)
(Tian et al., 2020)	Peking, Čína (N = 262)	20. 1. - 10. 2. 2020	5 %

* studie jsou v recenzním řízení