# Utilizing Functioning Data for 360 Degree Research: Using the Example of Prospective Budgeting and Performance-oriented Incentive Mechanism for Insurance Payment for Rehabilitation Episodes based on WHO's International Classification of Functioning, Disability and Health

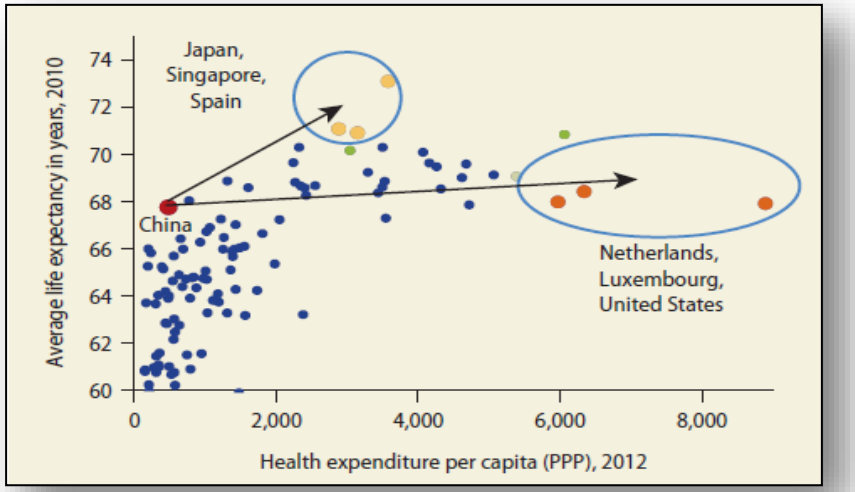
Jan D. Reinhardt, Sichuan University & Yuewei Ling, Stanford University

16 Feb 2024

# Starting point: Health expenditure trap

- Many OECD countries face growing health care spending with only marginal improvement in outcomes
- China needs to avoid this trap
- Optimizing health and minimizing costs is not possible at the same time
- Optimizing outcomes while controlling costs
- Therefore, better organization, management, and payment systems

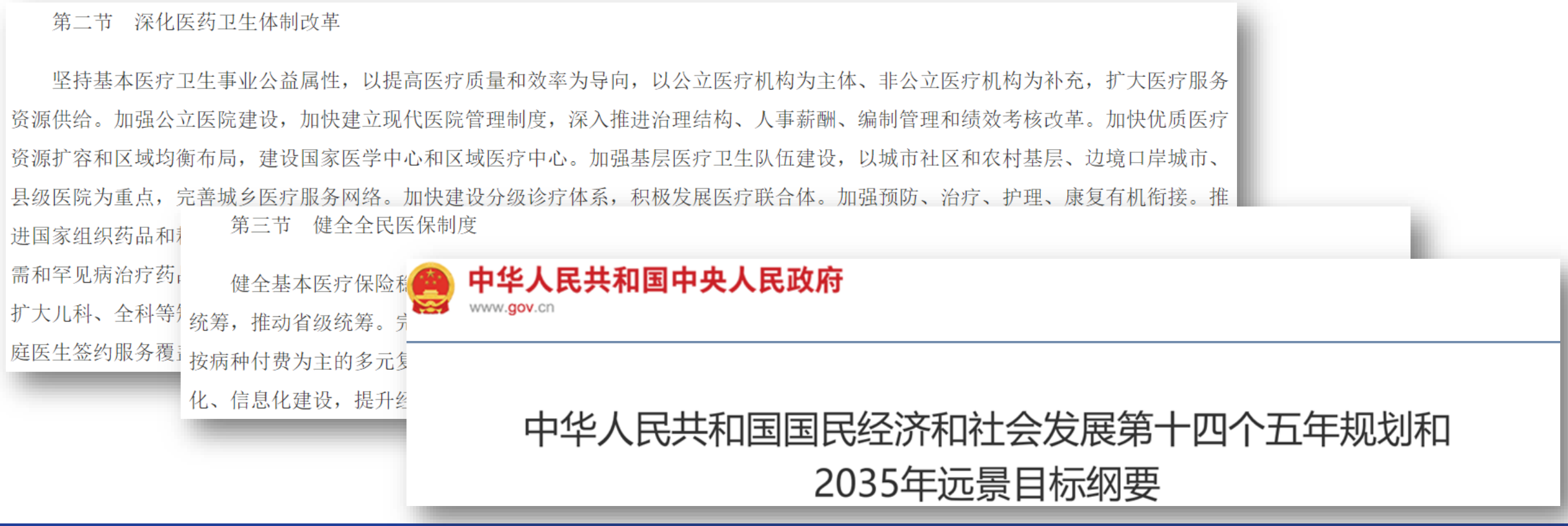
Life expectancy of selected countries by PPP health expenditure/head




Source: World Bank & WHO 2019: Healthy China

# What to do?

- 14th Five-year Plan of the PRC (2021-2025)
- Move from fee for services to prospective payment systems such as DRG/DIP
  - At same time, orientation to improve medical quality and efficiency



# However

- **Improvement in health not addressed in DRG and other payment approaches**
  "The various payment mechanisms provide conflicting incentives for 'productivity' and 'expenditure control' and **none provides positive incentives for quality**. As a result, many countries have blended payment mechanisms and/or have introduced reimbursements linked to desired outcomes (e.g. value-based payments) or models of care (e.g. integrated care). "
  (European Observatory on Health Systems and Policies, URL: https://eurohealthobservatory.who.int/themes/health-system-functions/financing-and-payment/payment-mechanisms, accessed 11 Feb 2024)
- **Cost estimation is mostly arrived at based on data from existing system**
  - **Estimated independent of patient improvement**

# Broken lamp example

- You got an expensive designer lamp which is broken.
  - Option A: Mr. Wang from the community offers to try to repair it for ￥300, it will likely take 3 hours and the lamp may not work.
  - Option B: Mr. Yang from a specialized company offers to do it for ￥500 to be paid upon success.
- How do you budget?
- If you choose Mr. Wang for ￥300, what will the final cost be?

# Application to health care

- Reduced cost option may not lead to improvement of health
- Greater costs may indeed occur elsewhere in the system (follow up treatment episode, social costs)

# However

- **Improvement in health not addressed.**
  "The various payment mechanisms provide conflicting incentives for 'productivity' and 'expenditure control' and **none provides positive incentives for quality**. As a result, many countries have blended payment mechanisms and/or have introduced reimbursements linked to desired outcomes (e.g. value-based payments) or models of care (e.g. integrated care). "
  (European Observatory on Health Systems and Policies, URL: https://eurohealthobservatory.who.int/themes/health-system-functions/financing-and-payment/payment-mechanisms, accessed 11 Feb 2024)
- **Cost estimation is mostly arrived at based on data from existing system**
  - **Estimated independent of patient improvement**

Our approach

➢ **Integrate functioning as core indicator of baseline health, treatment targeting, and outcome**

➢ **Create payment systems for health improvement**

# Objective and specific aims of the project

**Develop model of a prospective and performance-oriented payment system for rehabilitation based on baseline functioning and functioning improvement**

1. Establish ICF-based metric for functioning assessment/evaluation
2. Establish Minimal Important Difference (MID)
3. Demonstrate that (medical) costs are influenced by degree of improvement
4. Demonstrate that costs for improvement are associated with baseline function
5. Create model for functioning-related budget estimation within diagnostic group (here selected stroke as complex condition that is supposed to have a major effect on functioning)
6. Develop performance-oriented payment algorithm for functioning improvement over rehab period and assess needed investment and impact based on simulation models

# Information on data and analysis

- Any data shown are, unless otherwise specified, from an ongoing project on ICF-based insurance payment in China
- **Design and setting: Multi-center prospective inpatient cohort study involving 11 level III hospitals in Wuxi and Nanjing, Jiangsu, China**
- **Population: Here selected 2400 people with stroke (ischemic or hemorrhagic) admitted to rehab department of partner hospitals between 08/19/2020 and 06/26/2023**
  - Excluded: Comorbidity with disease in terminal stage, treatment target other than functioning (e.g. palliation, complications), mental disorder or substance abuse preventing collaboration with rehabilitation interventions
- Interventions: All patients received multi-disciplinary inpatient rehabilitation according to decisions of team in charge, spending not restricted to predefined services but based on decision of the leading physician (LJ)
- Comparator: None
- Outcome: Here total inpatient costs
- Analysis shown here is preliminary

# Establish assessment metric

- **ICF tool consisting of 17 categories selected from rehab set through "expert" conference**
- Clinical meaningful description of categories
- Clinician- or patient-rated with 0-10 NRS (no to complete problem)
- Scoring based on Rasch metric (planned)
- Here: Use simple sum score (for demonstration purposes), high score means more problems in functioning

**Body functions**

Mental functions – Global mental functions
- **b130 energy and drive**
- b134 sleep functions
- **b152 emotional functions**

Sensory functions and pain
- **b280 sensation of pain**

Genitourinary and reproductive functions
- b620 urination functions

Neuromusculoskeletal and movement-related functions
- b710 mobility of joint functions
- b730 muscle power functions

**Activity and Participation**

General tasks and demands
- **d230 carrying out daily routine**

Mobility
- d410 changing basic body position
- **d450 walking**
- d455 moving around
- **d465 moving around using equipment**

Self-care
- d510 washing oneself
- d530 toileting
- d540 dressing
- d550 eating

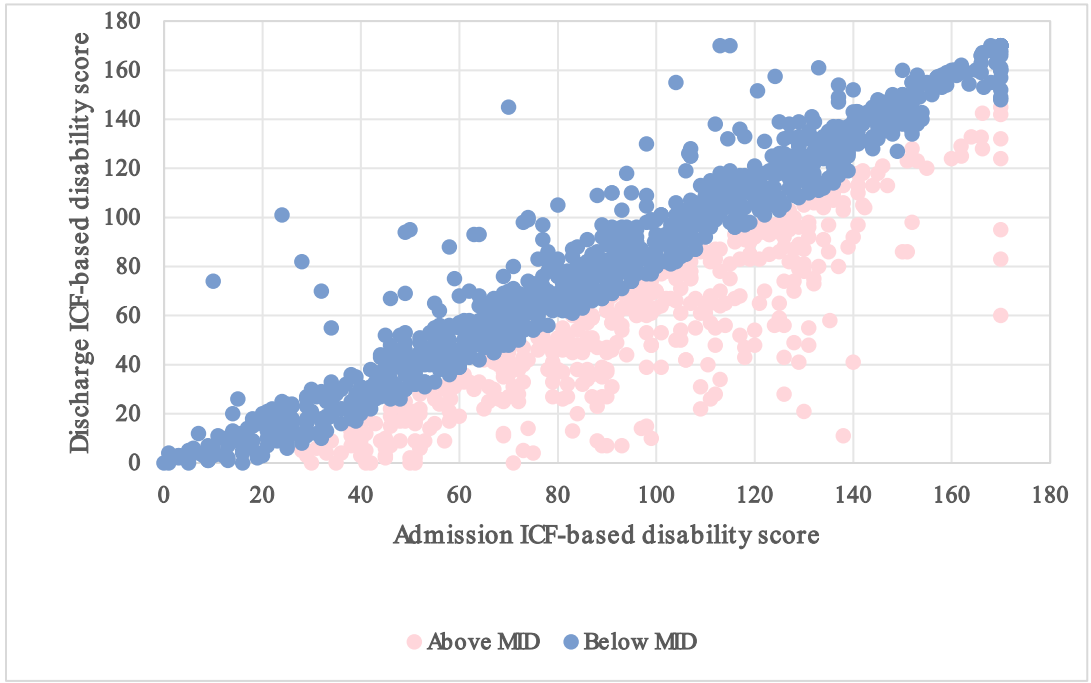
Interpersonal interactions/relationships
- d710 basic interpersonal interactions

**Generic-6 categories are in bold**

# Establish Minimal Important Difference (MID)

- Necessary to differentiate at least minimally improved population
- Then estimate costs only for above MID subpopulation to avoid underestimation
- Anchor-based using Barthel-Index improvement in one dependence level

- Arrived at MID of **-22.364 ± 1.801** (95% CI) for ICF-tool sum score
- Of 2020 patients **615** patients with ICF-tool sum score above MID

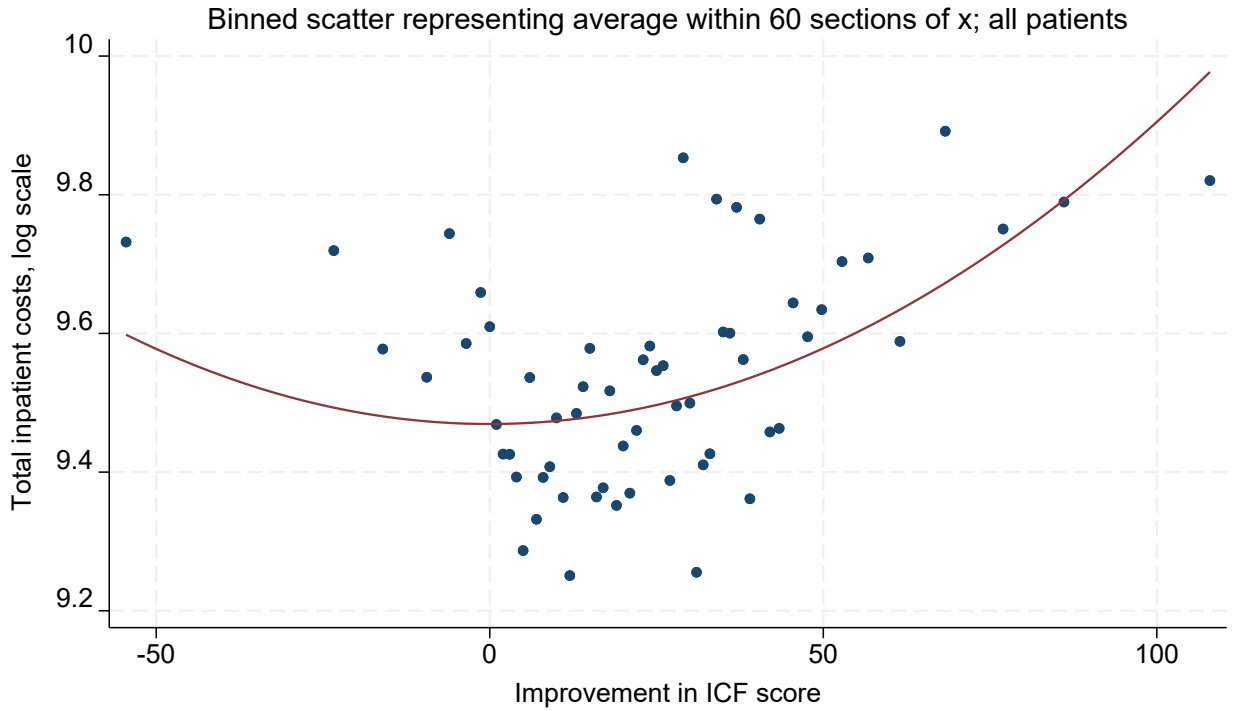


# Demonstrate that (medical) costs are influenced by degree of improvement

- Agame changer that can connect functioning improvement and prospective budget estimation

Distribution of total inpatient rehab costs in CNY by achievement of MID for ICF-based assessment tool

| | Above MID | Below MID |
|---|---|---|
| Number of Patients (total 2020) | 615 | 1405 |
| Hospitalization cost: **Minimum** | ¥ 1,826.0 | ¥ 826.0 |
| Hospitalization cost: **Maximum** | ¥ 286,232.0 | ¥ 211,157.0 |
| Hospitalization cost: **Mean ± SD** | ¥ 19,554.2 ± 22,925.8 | ¥ 16,598.8 ± 18,146.8 |



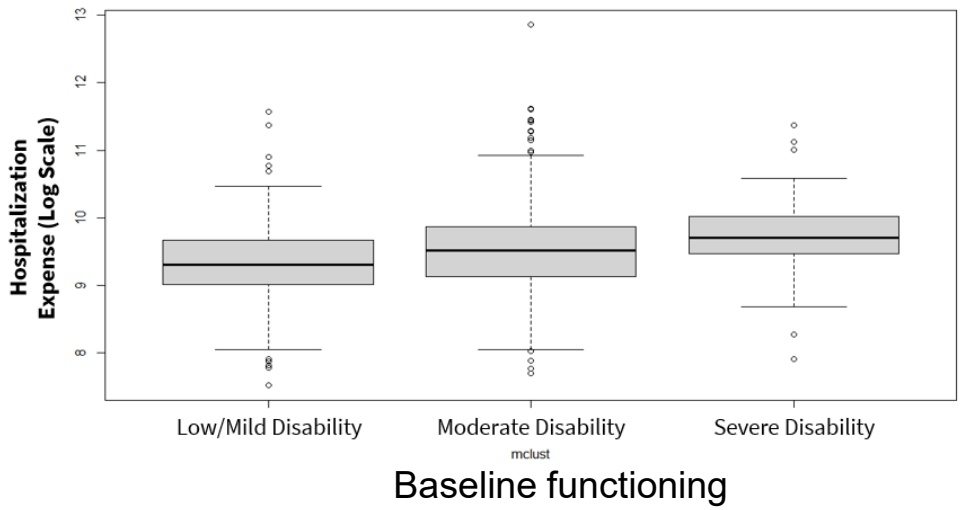

P for improvement in ICF score = 0.001 in mixed-effects Gamma regression with random intercept for hospital, adjusted for baseline ICF score, age, gender, education

# Demonstrate that costs for improvement are associated with baseline function

- Using k-prototypes clustering and MClust algorithms for grouping patients according to baseline function
- Based on 17 ICF categories
- Used for stratification of prediction model by patient groups (3 strata representing different baseline function) and group determination for benchmarking

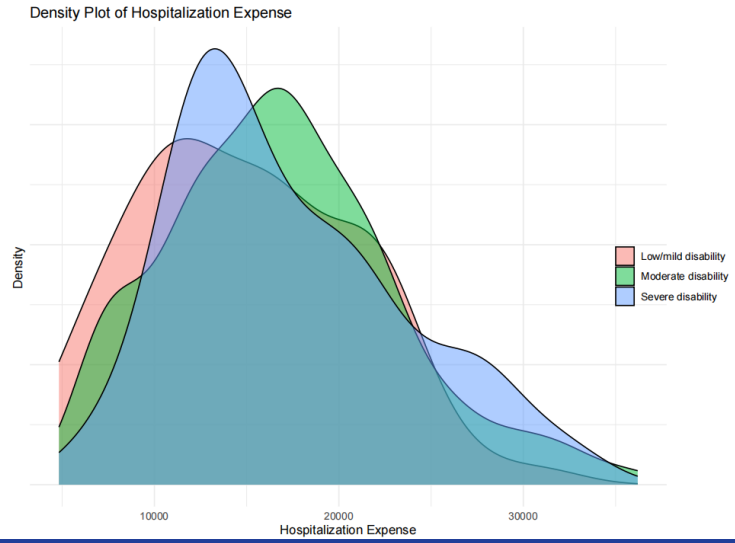
Total inpatient rehab costs in CNY by patient baseline functional status

| | Low/Mild Disability | Moderate Disability | Severe Disability |
|---|---|---|---|
| Average | 13,268 | 16,789 | 20,158 |
| First Quartile | 8,159 | 9,221 | 12,851 |
| Median | 10,907 | 13,553 | 16,276 |
| Third Quartile | 15,681 | 19,165 | 22,506 |



# Create model for functioning-related budget estimation

- **Adopted 12 machine learning algorithms to develop cost prediction model for prospective budgeting for population above MID**
  - Training to test data split randomly 70 to 30 percent
  - based on clinical information at admission, including ICF tool score, brain injury type, number of complications
  - adjusted by demographic data at admission, including age, gender, education level
- Support vector machine (SVM) performs best with RMSE values of 0.74 (excellent) on the training set, and 1.27 (good) on the test set
- Predicted mean hospitalization costs:
  - Low/mild disability: ￥13,398
  - Moderate disability: ￥15,234
  - Severe disability: ￥19,753
- For each incoming individual patient a case-tailored estimation can be arrived at



# Develop performance-oriented payment algorithm for functioning improvement over rehab period and assess needed investment and impact based on simulation models

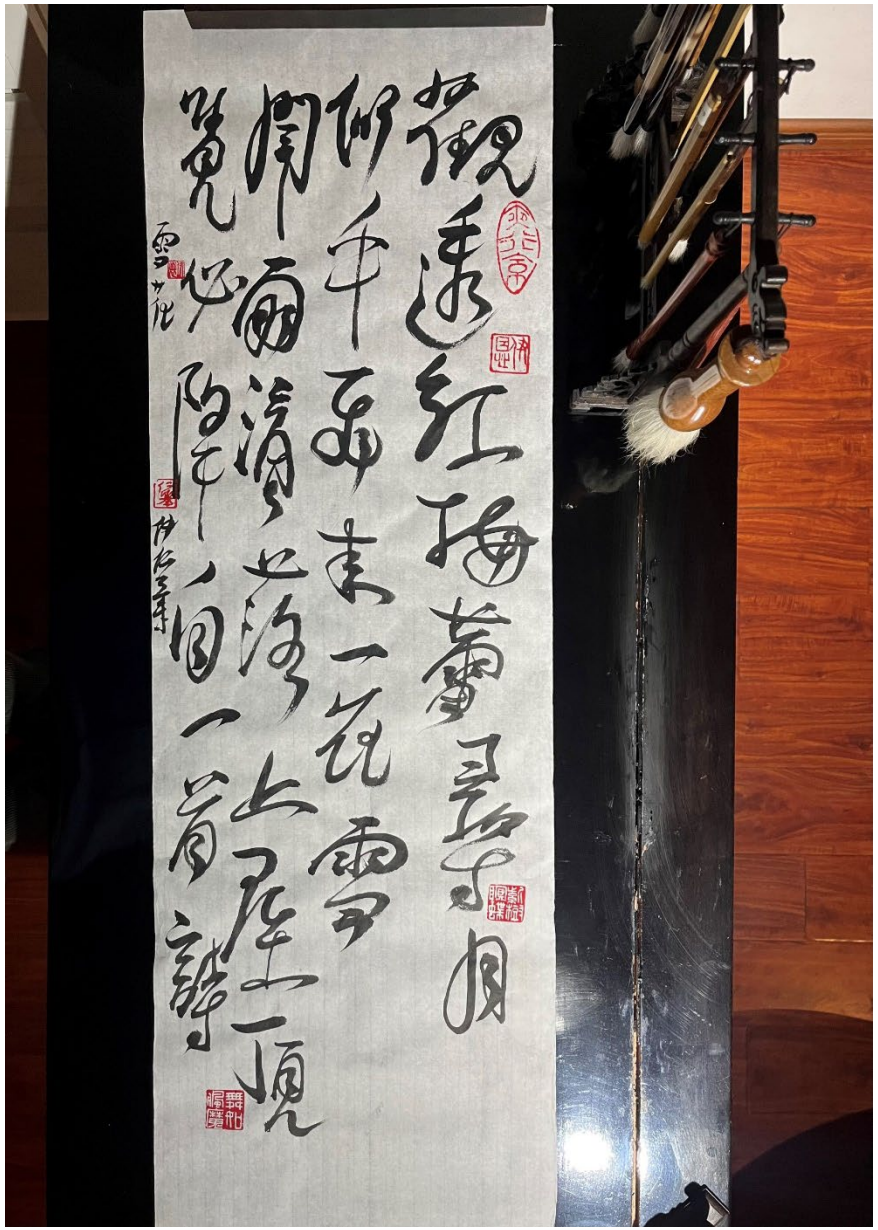
- Some ideas
  - For each group (above MID population), average ICF-based assessment score improvement adjusted for baseline function will be estimated with CIs (70-95%).
  - Upper bound of ICF-improvement CI (size to be determined) can serve as benchmark
  - Degree of performing above or below benchmark can be determined (based on total population)
  - Bonuses and rewards for average or individual patient-based performance to be set by insurances
  - Combination with other performance indicators possible

# Limitations of the ongoing study

- Approach applicable to health care that targets health improvement, not to health care targeting prevention, disease control (avoid deterioration), palliation
  - But can be adapted
- We use one MID for all patients, but dependent on baseline function, different MIDs may be determined
  - For example, somebody with high functioning at admission, can only be improved up to the scale ceiling.
  - Can be solved through adjustment for baseline function in ANCOVA style model
- Biased assessment through clinicians providing services (under-estimation of baseline, over-estimation of improvement)
  - Additional risk of data manipulation when used for performance evaluation
- Link to population health and overall hospital-based information systems to be established

# Conclusions

- Meaningful improvement of functioning needs larger investment than overall average
- Baseline functioning is an important predictor of costs within a diagnostic group
- Key performance indicators drawing on functioning information can be used for benchmarking and performance-oriented payment component
- Related research involves different approaches and tools and a series of methodological challenges needs to be addressed



Snowflake

Looking through the red mume's buds searching for the moon,

It seems to me as if there flew a little flake of snow.

Hearing yet the raindrops fall upon the houses' roofs,

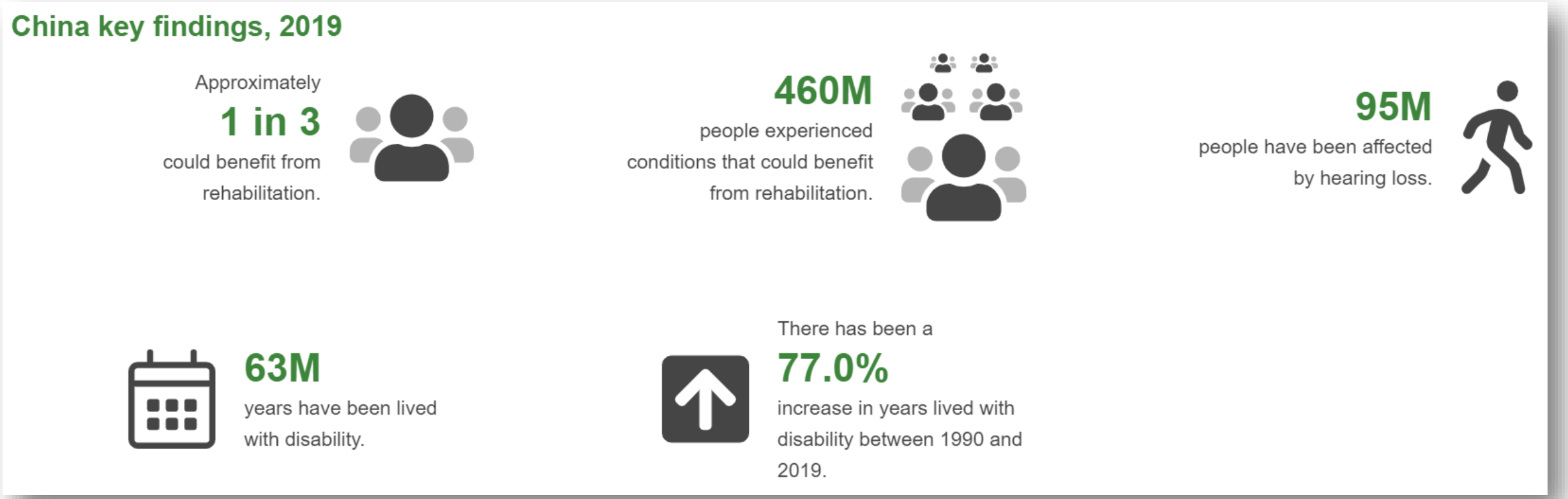
I feel it must have rather fallen out of an ancient poem.

# Acknowledgements and Funding


- We would like to express our gratitude to the following individuals: Prof. Jian'an Li for his invaluable guidance and unwavering friendship; Prof. Xuanxuan Wang, Dr. Shouguo Liu, Dr. Jiahui Li, Dr. Xia Zhang, Ms. Yangting Zeng, and Mr. Kuofu Liu for their exceptional project support.
- This project is funded by the National Ministry of Education of the People's Republic of China (23YJA63000)

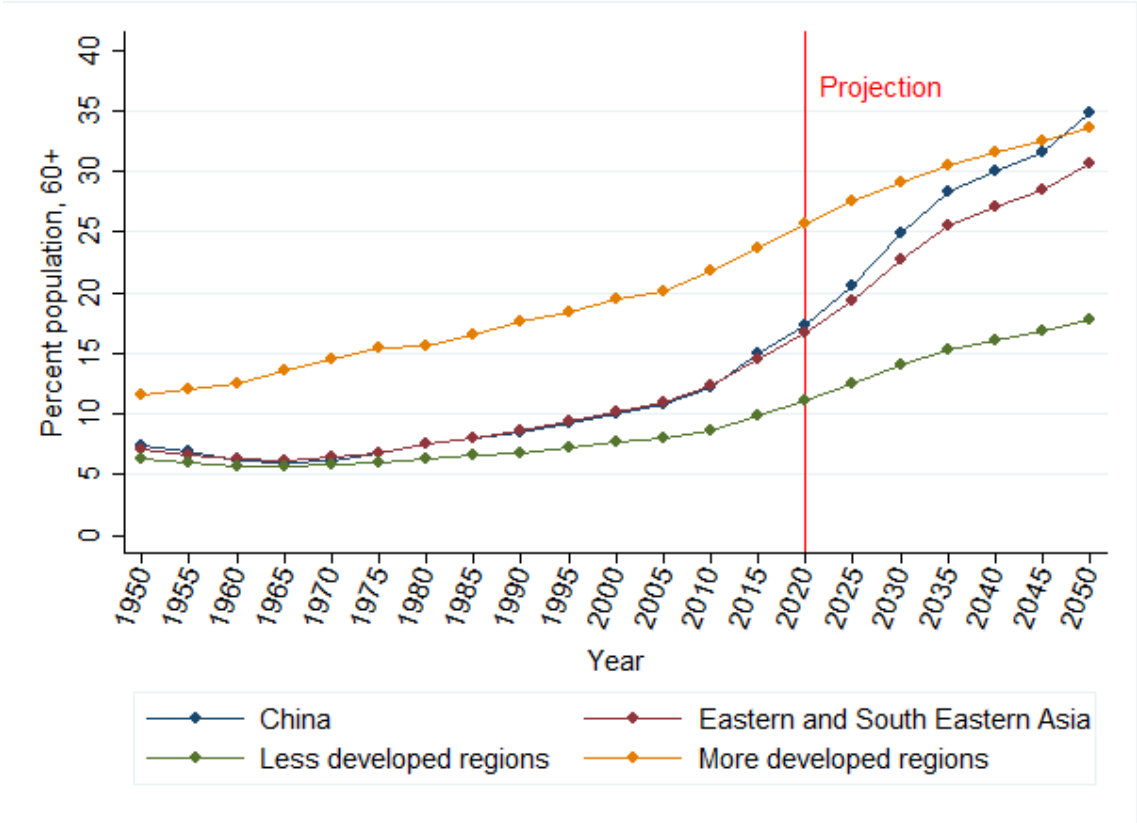
# Backup

# Rehabilitation needs in China

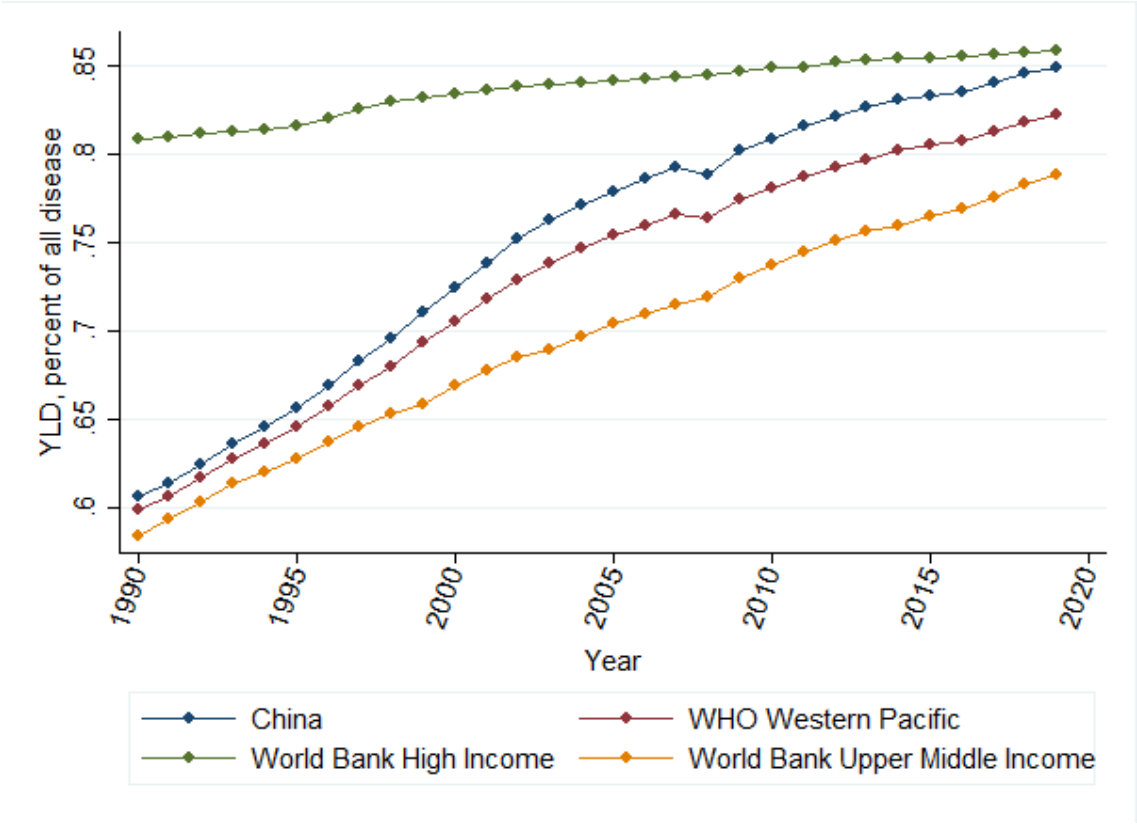



Source: https://vizhub.healthdata.org/rehabilitation, accessed April 25, 2021

## Demographic ageing in China, compared with selected countries




## Share of non-communicable disease



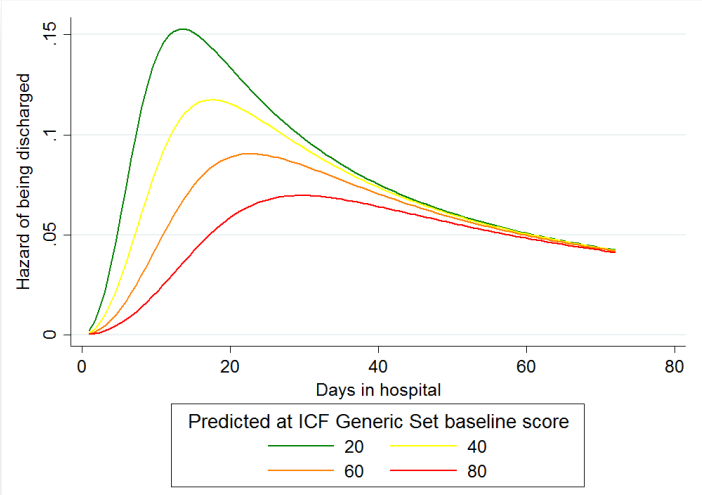
# Current status

- **Only one study in China demonstrated effect of baseline functioning (ICF-Generic Set) on length of stay and costs**
- Several Chinese model systems implement prospective budgeting based on DRG/DIP
  - Inconsistent results
  - Mostly cost saving, but sometimes opposite effects for patient sub-populations
  - **Effects on health outcomes not evaluated**
- Cost estimation is mostly arrived at based on data from existing system
  - Restricted list of services paid by insurance
  - **Estimated independent of patient improvement**
- Only few studies on performance-oriented payment components
  - Mostly understood in terms of adjustment factor
  - Reward on hospital, not team level
  - **Rarely integrated with prospective budgeting**

# Functioning information is relevant to resource utilization across departments

- 4510 Chinese (mainland) patients
- 50 hospitals, 20 provinces
- Departments of rehabilitation, orthopaedics, neurology, cardiology, pneumology, and cerebral surgery
- 6 diagnostic groups

Gamma shared frailty loglogistic survival regression of discharge probability on baseline functioning, adjusted for diagnosis, department, age, gender




ORIGINAL RESEARCH | VOLUME 100, ISSUE 8, P1450-1457.E1, AUGUST 01, 2019

Purchase Subscribe

## System-wide Clinical Assessment of Functioning Based on the International Classification of Functioning, Disability and Health in China: Interrater Reliability, Convergent, Known Group, and Predictive Validity of the ICF Generic-6

Shouguo Liu, MD * • Jan D. Reinhardt, PhD * • Xia Zhang, PhD • ... Birgit Prodinger, PhD • Shan Liu, MSc • Jianan Li, MD • Show all authors • Show footnotes

Published: December 14, 2018 • DOI: https://doi.org/10.1016/j.apmr.2018.11.014 • Check for updates

**Table 4** Coefficients from gamma shared frailty loglogistic survival model for length of hospital stay (time ratios) and from mixed effects generalized linear model of gamma family for costs of treatment

| | Length of Hospital Stay (d) (n=4508) | | | | Treatment Costs in Chinese Yuan (n=4293) | | | |
|---|---|---|---|---|---|---|---|---|
| Predictor | Time Ratio | 95% CI | | *P* Value | Coefficient | 95% CI | | *P* Value |
| ICF GS 6 at baseline | 1.013 | 1.011 | 1.015 | <.001 | 0.009 | 0.007 | 0.011 | <.001 |
| Length of stay (d) | NA | NA | NA | NA | 0.044 | 0.041 | 0.047 | <.001 |
| Diagnostic group (reference: musculoskeletal) | | | | | | | | |
| Cancer | 1.294 | 1.130 | 1.481 | <.001 | −0.047 | −0.198 | 0.105 | .546 |
| Cardiovascular | 1.103 | 0.957 | 1.271 | .176 | 0.191 | 0.014 | 0.369 | .034 |
| Respiratory | 1.238 | 1.077 | 1.423 | .003 | −0.081 | −0.243 | 0.082 | .33 |
| Neurologic | 1.402 | 1.293 | 1.521 | <.001 | 0.044 | −0.050 | 0.138 | .361 |
| Others | 1.223 | 1.364 | 0.311 | <.001 | −0.022 | −0.145 | 0.101 | .725 |
| Hospital department | | | | | | | | |
| Neurology | 0.605 | 0.566 | 0.646 | <.001 | 0.192 | 0.111 | 0.273 | <.001 |
| Cerebral surgery | 0.738 | 0.671 | 0.812 | <.001 | 0.970 | 0.869 | 1.071 | <.001 |
| Orthopedics | 0.888 | 0.820 | 0.961 | .003 | 0.955 | 0.866 | 1.044 | <.001 |
| Pneumology | 0.538 | 0.475 | 0.609 | <.001 | 0.285 | 0.139 | 0.431 | <.001 |
| Cardiology | 0.550 | 0.483 | 0.626 | <.001 | 0.637 | 0.474 | 0.800 | <.001 |
| Intercept | 9.721 | 8.732 | 10.822 | <.001 | 8.234 | 8.071 | 8.398 | <.001 |
| Loglikelihood | −3529.688 | | | | −45618.905 | | | |

Abbreviations: ICF GS 6, ICF Generic 6 Rasch-transformed interval scale score; NA, not applicable.

# Current status

- Only one study in China demonstrated effect of baseline functioning (ICF-Generic Set) on length of stay and costs
- Several Chinese model systems implement prospective budgeting based on DRG/DIP
  - Inconsistent results
  - Mostly cost saving, but sometimes opposite effects for patient sub-populations
  - Effects on health outcomes not evaluated
- **Cost estimation is mostly arrived at based on data from existing system**
  - Restricted list of services paid by insurance
  - **Estimated independent of patient improvement**
- Only few studies on performance-oriented payment components
  - Mostly understood in terms of adjustment factor
  - Reward on hospital, not team level
  - Rarely integrated with prospective budgeting

# Current status

- Only one study in China demonstrated effect of baseline functioning (ICF-Generic Set) on length of stay and costs
- Several Chinese model systems implement prospective budgeting based on DRG/DIP
  - Inconsistent results
  - Mostly cost saving, but sometimes opposite effects for patient sub-populations
  - Effects on health outcomes not evaluated
- Cost estimation is mostly arrived at based on data from existing system
  - Restricted list of services paid by insurance
  - Estimated independent of patient improvement
- Only few studies on performance-oriented payment components
  - Mostly understood in terms of adjustment factor
  - Reward on hospital, not team level
  - **Rarely integrated with prospective budgeting**

# Idea: 6-point rationale

1. Main target of rehabilitation and other health care is not diagnosis (which is just a label) but health
2. Functioning as health outcome and health system input (at admission) is WHO's operational definition of health and can be classified with WHO's International Classification of Functioning, Disability and Health (ICF)
3. Functioning improvement should be a key feature for health services payment systems.
4. Budget estimation based on baseline function can inform functioning related algorithms or grouping within DRG to more accurately predict needed financial input and at the same time create leeway for decision making of medical specialists.
5. Functioning improvement as a core health outcome should be the basis for such prospective budget estimation, that is the estimates should be based on a population that has meaningfully improved.
6. Functioning improvement should be a key performance indicator that is used for benchmarking and performance-oriented payment to provide incentives for the continuous improvement of health care towards its target.

# Establish metric and assessment: Challenges and open questions

- Selection of the 17 categories questionable
- Differential representation of dimensions such as 4 items for mobility, self-care
  - Possible multi-dimensional Rasch
- Differential impact on costs
  - Possible feature selection or weighting for cost estimation
  - Working with profiles instead of scores
- Rehabilitation targeting/goal setting not considered
- **Biased assessment through clinicians providing services (under-estimation of baseline, over-estimation of improvement)**

# Establish MID: Challenges and open questions

- Dependence on baseline functioning to be considered
  - Use of ANCOVA style model
  - Stratify by baseline functioning
- Ceiling effect of Barthel Index as anchor
  - Combine with standardized effect size approach
  - Combine with other anchors (particularly, patient-reported outcome measures)
- **Non-binary/continuous approach to be considered as greater improvement than MID may be desirable**

# Degree of improvement influence on costs: Challenges &open questions

- Differentiation between types of costs such as administrative vs. medical
- Differentiation of medical costs into pharma, therapy, diagnostic, etc.
- Consideration of comorbidity (type and number of conditions)
- Considering co-interventions running parallel to rehab
- Dealing with outliers

# Budget estimation based on baseline function: Challenges &open questions

- Functioning score vs profile?
- Differential influence of ICF categories on costs
- Differentiation between cost types
- Rehabilitation targeting not considered
- Overall medical costs such as re-hospitalization costs and social costs not considered

# Performance-oriented payment algorithm: Challenges and open questions

- How to establish simulation models and which data to feed them?
- How to consider inpatient mortality in estimation of average improvement of department patient population?
- How to avoid biases and tendencies by departments/teams to manipulate data?
- How to finance the incentive pool?

# Strength of present study

- Estimation under conditions of freedom of clinical decision making (black box model)
  - Normally, cost estimation is based on old healthcare system
  - without freedom of applying treatments based on clinical decisions
  - Doctors have to select from insurance lists
  - In our model system, insurance will pay for the treatment independently of their predefined services list
- Estimations based on patients who actually improve
- Multi-center study with homogeneous population
- Up-to-date machine learning prediction models linked with cluster analysis and MID determination