

OHDSI内では、実名での活動になります。
Zoom参加時も「名前は実氏名で」お願いします。

OHDSI Japan evening conference #21

イブニング カンファレンス(第21回)

2021.8.25



OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

オデッセイ
ジャパン



2021 OHDSI global symposium

<https://www.ohdsi.org/2021-ohdsi-global-symposium-info/>

- 毎年開催のOHDSIシンポジウム
- 今年もWeb開催(MS-Teams)
- 日程: 2021年9月12-15日

上記URLから登録受付中！

参加費無料、寄付は受付けてます。



APAC Community Callのお知らせ

- 隔週で、木曜日の日本時間お昼に開催中
- 9月はglobal symposiumがあるため、次回開催日程は下記を確認してください。

OHDSI APACページ <https://ohdsi.org/apac/>

“[by using this direct link.](#)” からゲスト参加可能。

その前のリンクを使って、Teamsへの登録を推奨。

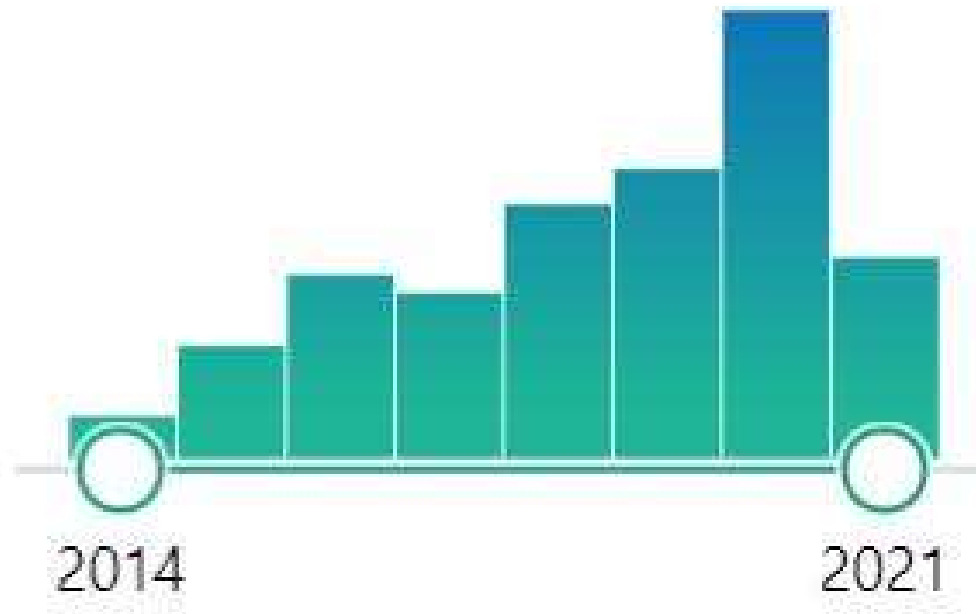
STEP1) [request access to our MS Teams Environment](#)

STEP2) request access to [our OHDSI APAC workgroup](#)



OHDSI論文

Pubmedで"OHDSI"を検索



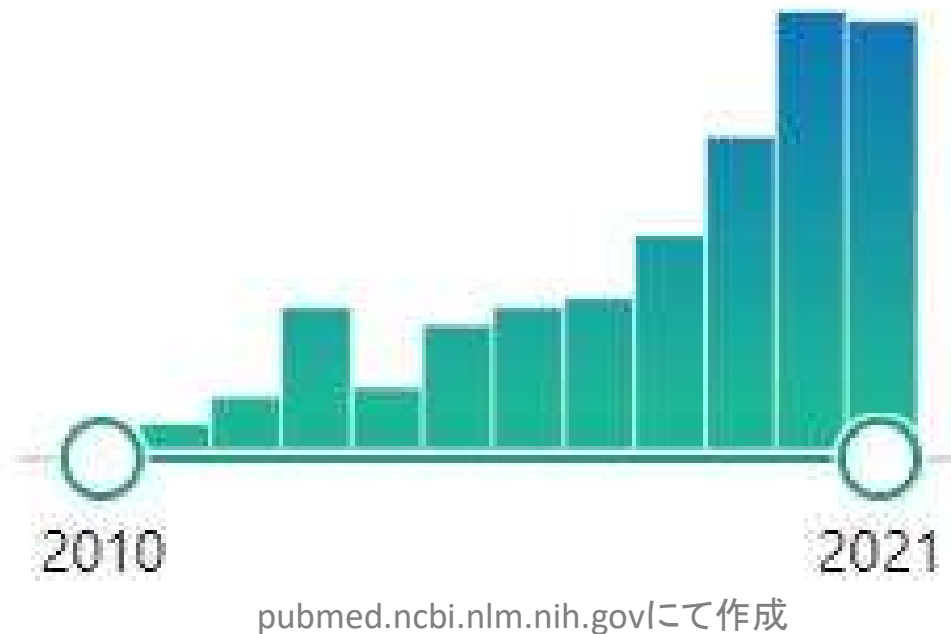
pubmed.ncbi.nlm.nih.govにて作成

累計：7月91本 → 8月92本



OHDSI関連論文（今月から変更）

Pubmedで"OHDSI or OMOP"を検索



累計：8月210本



➤ [Appl Clin Inform.](#) 2021 Aug;12(4):757-767. doi: 10.1055/s-0041-1732301. Epub 2021 Aug 11.

Transformation of Electronic Health Records and Questionnaire Data to OMOP CDM: A Feasibility Study Using SG_T2DM Dataset

Selva Muthu Kumaran Sathappan¹, Young Seok Jeon¹, Trung Kien Dang¹, Su Chi Lim², Yi-Ming Shao², E Shyong Tai³, Mengling Feng^{1 4}

Affiliations — collapse

Affiliations

- 1 Saw Swee Hock School of Public Health, National University Health System and National University of Singapore, Singapore, Singapore.
- 2 Clinical Research Unit, Khoo Teck Puat Hospital, Singapore, Singapore.
- 3 Division of Endocrinology, National University Hospital, Singapore, Singapore.
- 4 Institute of Data Science, National University of Singapore, Singapore, Singapore.

PMID: 34380168 PMCID: [PMC8357458](#) DOI: [10.1055/s-0041-1732301](#)

Abstract

Background: Diabetes mellitus (DM) is an important public health concern in Singapore and places a massive burden on health care spending. Tackling chronic diseases such as DM requires innovative strategies to integrate patients' data from diverse sources and use scientific discovery to inform clinical practice that can help better manage the disease. The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) was chosen as the framework for integrating data with disparate formats.

Objective: The study aimed to evaluate the feasibility of converting Singapore based data source, comprising of electronic health records (EHR), cognitive and depression assessment questionnaire data to OMOP CDM standard. Additionally, we also validate whether our OMOP CDM instance is fit for the purpose of research by executing a simple treatment pathways study using Atlas, a graphical user interface tool to conduct analysis on OMOP CDM data as a proof of concept.

Methods: We used de-identified EHR, cognitive, and depression assessment questionnaires data from a tertiary care hospital in Singapore to convert it to version 5.3.1 of OMOP CDM standard. We evaluate the OMOP CDM conversion by (1) assessing the mapping coverage (that is the percentage of source terms mapped to OMOP CDM standard); (2) local raw dataset versus CDM dataset analysis; and (3) Implementing Harmonized Intrinsic Data Quality Framework using an open-source R package called Data Quality Dashboard.

Results: The content coverage of OMOP CDM vocabularies is more than 90% for clinical data, but only around 11% for questionnaire data. The comparison of characteristics between source and target data returned consistent results and our transformed data did not pass 38 (1.4%) out of 2,622 quality checks.

Conclusion: Adoption of OMOP CDM at our site demonstrated that EHR data are feasible for standardization with minimal information loss, whereas challenges remain for standardizing cognitive and depression assessment questionnaire data that requires further work.

概要

背景: 糖尿病(DM)は、シンガポールにおける重要な公衆衛生上の問題であり、医療費に多大な負担をかけています。糖尿病のような慢性疾患に対処するためには、様々な情報源から得られる患者のデータを統合し、科学的発見を利用して、疾患の管理を改善するための臨床実践を行う革新的な戦略が必要です。そこで、異なるフォーマットのデータを統合するためのフレームワークとして、OMOP(Observational Medical Outcomes Partnership)のCDM(Common Data Model)を採用しました。

目的: 本研究の目的は、シンガポールを拠点とする電子カルテ(EHR)、認知・抑うつ評価アンケートデータからなるデータソースをOMOP CDM規格に変換することの実現性を評価することです。また、概念実証として、OMOP CDMデータの分析を行うためのグラフィカルユーザーインターフェースツールAtlasを用いて、簡単な治療パスウェイ研究を実施することで、OMOP CDMインスタンスが研究目的に適しているかどうかを検証しました。

方法: シンガポールの三次医療機関から得られた、非識別化されたEHR、認知機能、うつ病評価の質問票データを用いて、OMOP CDM標準のバージョン5.3.1に変換した。OMOP CDMへの変換を評価するために、(1)マッピングカバレッジ(OMOP CDM標準にマッピングされたソースタームの割合)の評価、(2)ローカルの生データセットとCDMデータセットの分析、(3)Data Quality Dashboardと呼ばれるオープンソースのRパッケージを用いたHarmonized Intrinsic Data Quality Frameworkの実装、を行った。

結果: OMOP CDMポキャブラリのコンテンツカバレッジは、臨床データでは90%以上であるが、アンケートデータでは11%程度である。ソースデータとターゲットデータの特性を比較したところ、一貫した結果が得られ、変換したデータは2,622件の品質チェックのうち38件(1.4%)を通過しなかった。

結論: 当院でのOMOP CDMの採用により、EHRデータは情報の損失を最小限に抑えて標準化が可能であることが実証されたが、認知機能とうつ病の評価のための質問票データの標準化には課題が残り、今後の検討が必要である。



➤ [J Am Med Inform Assoc. 2021 Jul 27;ocab132. doi: 10.1093/jamia/ocab132. Online ahead of print.](#)

Increasing trust in real-world evidence through evaluation of observational data quality

Clair Blacketer ^{1 2}, Frank J Defalco ¹, Patrick B Ryan ^{1 3}, Peter R Rijnbeek ²

Affiliations — collapse

Affiliations

- ¹ Observational Health Data Analytics, Janssen Research and Development, LLC, Titusville, New Jersey, USA.
- ² Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands.
- ³ Department of Biomedical Informatics, Columbia University, New York, New York, USA.

PMID: 34313749 DOI: [10.1093/jamia/ocab132](#)

Abstract

Objective: Advances in standardization of observational healthcare data have enabled methodological breakthroughs, rapid global collaboration, and generation of real-world evidence to improve patient outcomes. Standardizations in data structure, such as use of common data models, need to be coupled with standardized approaches for data quality assessment. To ensure confidence in real-world evidence generated from the analysis of real-world data, one must first have confidence in the data itself.

Materials and methods: We describe the implementation of check types across a data quality framework of conformance, completeness, plausibility, with both verification and validation. We illustrate how data quality checks, paired with decision thresholds, can be configured to customize data quality reporting across a range of observational health data sources. We discuss how data quality reporting can become part of the overall real-world evidence generation and dissemination process to promote transparency and build confidence in the resulting output.

Results: The Data Quality Dashboard is an open-source R package that reports potential quality issues in an OMOP CDM instance through the systematic execution and summarization of over 3300 configurable data quality checks.

Discussion: Transparently communicating how well common data model-standardized databases adhere to a set of quality measures adds a crucial piece that is currently missing from observational research.

Conclusion: Assessing and improving the quality of our data will inherently improve the quality of the evidence we generate.

概要

目的: 医療観察データの標準化が進んだことにより、方法論の飛躍的進歩、迅速な国際共同研究、そして患者の転帰を改善するための現実的なエビデンスの創出が可能となった。共通のデータモデルの使用など、データ構造の標準化は、データ品質評価の標準化されたアプローチと結合する必要がある。実世界のデータを分析して得られた実世界のエビデンスを信頼するためには、まずデータそのものを信頼しなければならない。

材料と方法: 適合性、完全性、妥当性のデータ品質フレームワークにおけるチェックタイプの実装を、検証と妥当性の両面から説明する。データ品質チェックと判断しきい値を組み合わせることで、さまざまな観察健康データソースのデータ品質報告をカスタマイズする方法を説明する。データ品質報告は、透明性を促進し、結果として得られるアウトプットに対する信頼を構築するために、現実世界でのエビデンスの生成と普及のプロセス全体の一部となり得ることを説明する。

結果: データ品質ダッシュボードはオープンソースのRパッケージであり、3300以上の設定可能なデータ品質チェックを体系的に実行・要約することで、OMOP CDMインスタンスにおける潜在的な品質問題を報告する。

考察: 共通のデータモデルで標準化されたデータベースが、一連の品質基準にどの程度準拠しているかを透明に伝えることは、現在の観察研究に欠けている重要な要素を追加することになる。

結論: データの質を評価し改善することは、生成するエビデンスの質を本質的に改善することになる。



> JMIR Med Inform. 2021 Jul 21;9(7):e25531. doi: 10.2196/25531.

Evaluation of Three Feasibility Tools for Identifying Patient Data and Biospecimen Availability: Comparative Usability Study

Christina Schüttler ¹, Hans-Ulrich Prokosch ¹, Martin Sedlmayr ², Brita Sedlmayr ²

Affiliations — collapse

Affiliations

- 1 Chair of Medical Informatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany.
- 2 Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Dresden, Germany.

PMID: 34287211 DOI: 10.2196/25531

Abstract

Background: To meet the growing importance of real-world data analysis, clinical data and biosamples must be timely made available. Feasibility platforms are often the first contact point for determining the availability of such data for specific research questions. Therefore, a user-friendly interface should be provided to enable access to this information easily. The German Medical Informatics Initiative also aims to establish such a platform for its infrastructure. Although some of these platforms are actively used, their tools still have limitations. Consequently, the Medical Informatics Initiative consortium MIRACUM (Medical Informatics in Research and Care in University Medicine) committed itself to analyzing the pros and cons of existing solutions and to designing an optimized graphical feasibility user interface.

Objective: The aim of this study is to identify the system that is most user-friendly and thus forms the best basis for developing a harmonized tool. To achieve this goal, we carried out a comparative usability evaluation of existing tools used by researchers acting as end users.

Methods: The evaluation included three preselected search tools and was conducted as a qualitative exploratory study with a randomized design over a period of 6 weeks. The tools in question were the MIRACUM i2b2 (Informatics for Integrating Biology and the Bedside) feasibility platform, OHDSI's (Observational Health Data Sciences and Informatics) ATLAS, and the Sample Locator of the German Biobank Alliance. The evaluation was conducted in the form of a web-based usability test (usability walkthrough combined with a web-based questionnaire) with participants aged between 26 and 63 years who work as medical doctors.

Results: In total, 17 study participants evaluated the three tools. The overall evaluation of usability, which was based on the System Usability Scale, showed that the Sample Locator, with a mean System Usability Scale score of 77.03 (SD 20.62), was significantly superior to the other two tools (Wilcoxon test; Sample Locator vs i2b2: $P=.047$; Sample Locator vs ATLAS: $P=.001$). i2b2, with a score of 59.83 (SD 25.36), performed significantly better than ATLAS, which had a score of 27.81 (SD 21.79; Wilcoxon test; i2b2 vs ATLAS: $P=.005$). The analysis of the material generated by the usability walkthrough method confirmed these findings. ATLAS caused the most usability problems ($n=66$), followed by i2b2 ($n=48$) and the Sample Locator ($n=22$). Moreover, the Sample Locator achieved the highest ratings with respect to additional questions regarding satisfaction with the tools.

Conclusions: This study provides data to develop a suitable basis for the selection of a harmonized tool for feasibility studies via concrete evaluation and a comparison of the usability of three different types of query builders. The feedback obtained from the participants during the usability test made it possible to identify user problems and positive design aspects of the individual tools and compare them qualitatively.

概要

背景: リアルワールドデータ解析の重要性が高まる中、臨床データやバイオサンプルをタイムリーに利用できるようにする必要がある。実現可能性プラットフォームは、特定の研究課題に対するこのようなデータの利用可能性を判断するための最初のコンタクトポイントとなることが多い。そのため、これらの情報に簡単にアクセスできるよう、ユーザーフレンドリーなインターフェースを提供する必要がある。German Medical Informatics Initiativeも、その基盤となるこのようなプラットフォームの構築を目指している。これらのプラットフォームの中には、積極的に利用されているものもあるが、そのツールにはまだ限界がある。そこで、Medical Informatics InitiativeのコンソーシアムであるMIRACUM (Medical Informatics in Research and Care in University Medicine) は、既存のソリューションの長所と短所を分析し、最適なグラフィカル・フィジビリティ・ユーザー・インターフェースを設計することに取り組んだ。

目的: 本研究の目的は、最も使いやすいシステムを特定し、その結果、調和のとれたツールを開発するための最良の基盤を形成することである。この目的を達成するために、エンドユーザーとしての研究者が使用する既存のツールの比較ユーザビリティ評価を行った。

方法: 評価は、あらかじめ選択した3つの検索ツールを対象とし、6週間にわたり無作為化設計による質的探索研究として実施した。対象としたツールは、MIRACUM i2b2 (Informatics for Integrating Biology and the Bedside) の実現可能性プラットフォーム、OHDSI (Observational Health Data Sciences and Informatics) のATLAS、そして German Biobank AllianceのSample Locator。評価は、医師として働く26歳から63歳までの参加者を対象に、ウェブベースのユーザビリティテスト(ユーザビリティウォークスルーとウェブベースのアンケートを組み合わせたもの)の形で実施した。

結果: 合計で17人の研究参加者が3つのツールを評価した。System Usability Scaleに基づいたユーザビリティの総合評価では、Sample LocatorがSystem Usability Scaleの平均スコア77.03 (SD 20.62) で、他の2つのツールよりも有意に優れていることが示された(Wilcoxon test; Sample Locator vs i2b2: $P=.047$; Sample Locator vs ATLAS: $P=.001$)。59.83 (SD 25.36) のスコアを持つi2b2は、27.81のスコアを持つATLASよりも有意に優れていた (SD 21.79; Wilcoxon test; i2b2 vs ATLAS: $P=.005$)。ユーザビリティ・ウォークスルー法で作成された資料の分析では、これらの結果が確認された。ATLASが最も多くのユーザビリティ問題を引き起こし ($n=66$)、次いでi2b2 ($n=48$)、Sample Locator ($n=22$) となった。さらに、ツールの満足度に関する追加質問に関しては、Sample Locatorが最も高い評価を得た。

結論: 本研究では、3種類のクエリビルダの使い勝手を具体的に評価・比較することで、フィジビリティスタディのための統一されたツールを選択するための適切な基盤を構築するためのデータを提供した。また、ユーザビリティテストで得られた参加者からのフィードバックにより、個々のツールのユーザの問題点やデザイン上の良い点を特定し、それらを定性的に比較することができた。



➤ [Gastroenterology](#). 2021 Jul 18;S0016-5085(21)03244-3. doi: 10.1053/j.gastro.2021.07.010.
Online ahead of print.

Outcomes of SARS-CoV-2 Infection in Patients With Chronic Liver Disease and Cirrhosis: A National COVID Cohort Collaborative Study

Jin Ge ¹, Mark J Pletcher ², Jennifer C Lai ³, N3C Consortium

Collaborators, Affiliations — collapse

Collaborators

N3C Consortium: [Jeremy R Harper](#), [Christopher G Chute](#), [Melissa A Haendel](#)

Affiliations

- ¹ Division of Gastroenterology and Hepatology, Department of Medicine, University of California-San Francisco, San Francisco, California. Electronic address: jin.ge@ucsf.edu.
- ² Department of Epidemiology and Biostatistics, University of California-San Francisco, San Francisco, California; Division of General Internal Medicine, Department of Medicine, University of California-San Francisco, San Francisco, California.
- ³ Division of Gastroenterology and Hepatology, Department of Medicine, University of California-San Francisco, San Francisco, California.

Abstract

Background & aims: In patients with chronic liver disease (CLD) with or without cirrhosis, existing studies on the outcomes with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection have limited generalizability. We used the National COVID Cohort Collaborative (N3C), a harmonized electronic health record dataset of 6.4 million, to describe SARS-CoV-2 outcomes in patients with CLD and cirrhosis.

Methods: We identified all patients with CLD with or without cirrhosis who had SARS-CoV-2 testing in the N3C Data Enclave as of July 1, 2021. We used survival analyses to associate SARS-CoV-2 infection, presence of cirrhosis, and clinical factors with the primary outcome of 30-day mortality.

Results: We isolated 220,727 patients with CLD and SARS-CoV-2 test status: 128,864 (58%) were noncirrhosis/negative, 29,446 (13%) were noncirrhosis/positive, 53,476 (24%) were cirrhosis/negative, and 8941 (4%) were cirrhosis/positive patients. Thirty-day all-cause mortality rates were 3.9% in cirrhosis/negative and 8.9% in cirrhosis/positive patients. Compared to cirrhosis/negative patients, cirrhosis/positive patients had 2.38 times adjusted hazard of death at 30 days. Compared to noncirrhosis/positive patients, cirrhosis/positive patients had 3.31 times adjusted hazard of death at 30 days. In stratified analyses among patients with cirrhosis with increased age, obesity, and comorbid conditions (ie, diabetes, heart failure, and pulmonary disease), SARS-CoV-2 infection was associated with increased adjusted hazard of death.

Conclusions: In this study of approximately 221,000 nationally representative, diverse, and sex-balanced patients with CLD; we found SARS-CoV-2 infection in patients with cirrhosis was associated with 2.38 times mortality hazard, and the presence of cirrhosis among patients with CLD infected with SARS-CoV-2 was associated with 3.31 times mortality hazard. These results provide an additional impetus for increasing vaccination uptake and further research regarding immune responses to vaccines in patients with severe liver disease.

概要

背景と目的: 肝硬変を伴う、あるいは伴わない慢性肝疾患(CLD)患者において、重症急性呼吸器症候群コロナウイルス2(SARS-CoV-2)感染による転帰に関する既存の研究は、一般化に限界がある。我々は、640万人の電子カルテデータを統合したNational COVID Cohort Collaborative(N3C)を用いて、CLDおよび肝硬変の患者におけるSARS-CoV-2感染症の転帰について検討した。

方法: 2021年7月1日時点で、N3C Data EnclaveでSARS-CoV-2検査を受けた肝硬変を伴うまたは肝硬変を伴わないCLD患者をすべて特定した。生存分析を用いて、SARS-CoV-2感染、肝硬変の有無、臨床的要因を、主要アウトカムである30日死亡率と関連づけた。

結果: CLDおよびSARS-CoV-2検査のステータスを持つ220,727人の患者を分離した。12万8,864人(58%)が非肝硬変／陰性、2万9,446人(13%)が非肝硬変／陽性、5万3,476人(24%)が肝硬変／陰性、8941人(4%)が肝硬変／陽性であった。30日間の全死亡率は、肝硬変／陰性患者で3.9%、肝硬変／陽性患者で8.9%だった。肝硬変/陰性患者と比較して、肝硬変/陽性患者の30日後の死亡の調整ハザードは2.38倍であった。非肝硬変/陽性患者と比較して、肝硬変/陽性患者は30日後の死亡の調整ハザードが3.31倍であった。年齢、肥満、併存疾患(糖尿病、心不全、肺疾患)が増加した肝硬変患者の層別解析では、SARS-CoV-2感染は死亡の調整後ハザードの増加と関連していた。

結論: 本研究では、全国を代表する多様な性別のCLD患者約221,000人を対象とし、肝硬変患者のSARS-CoV-2感染は死亡ハザードの2.38倍、SARS-CoV-2に感染したCLD患者の肝硬変の有無は死亡ハザードの3.31倍と関連していることを明らかにした。これらの結果は、重度の肝疾患患者におけるワクチンの接種率を高め、ワクチンに対する免疫反応に関する研究をさらに進めるための新たな原動力となる。



OMOP環境の実装手順

疑似データをPostgres DBに入れる

今回使う疑似データ(合成データ)

SynPUFs: Medicare Claims Synthetic Public Use Files

Medicareは米国の公的医療保険。65歳以上が対象。

PostgreSQL: 今回はVersion 10を使用(最新版は13)

手順

1. PostgreSQLを導入
2. OMOP-CDM table定義を準備
3. SynPUFの入手、展開と加工
4. Tableへの読み込み